

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 18:12:45 ; Search time 7189.25 Seconds  
(without alignments)  
17088.240 Million cell updates/sec

Title: US-09-935-290-1  
Perfect score: 3003  
Sequence: 1 ttoggcaccaggtctgctgcg.....ctggagatcttcccagctttt 3003

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sv.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3001	99.9	3003	6	AX420453 Sequence
2	2958	98.5	3210	6	AX451419 Sequence
3	2755.4	91.8	4918	9	AL833093 Homo sapi
4	2752.2	91.6	3490	9	BC030783 Homo sapi
5	2752.2	91.6	6387	9	AL833061 Homo sapi
6	2485.4	82.8	2487	6	AX431391 Sequence
7	2484	82.7	2484	6	AX420455 Sequence
8	2241.8	74.7	5689	9	AL832464 Homo sapi
9	2131	71.0	5674	9	AB046780 Homo sapi
10	1917.6	63.9	3857	10	BC019201 Mus muscu
11	1903.8	63.4	6634	10	MUSG3PAT
12	1896.2	63.1	2646	6	AX401972 Sequence
13	1896.2	63.1	2646	10	AF021348 Rattus no
14	1807.2	60.2	2519	10	U36771 Rattus norv
15	267.4	8.9	196657	9	AL391986 Human DNA
16	246.4	8.2	271	6	BD029418 Sequence
17	223.6	7.4	255430	2	AC099387 Rattus no
18	218.2	7.3	197741	2	AC110206 Mus muscu
19	203.4	6.8	2000	6	AX431393 Sequence
20	173.8	5.8	206914	2	AC104911 Mus muscu
21	142.6	4.7	4201	4	AF469047 Bos tauru
22	131.6	4.4	147242	2	AC135152 Rattus no
23	123.2	4.1	147242	2	AC135152 Rattus no
24	109	3.6	2683	6	AX710670 Sequence
25	109	3.6	2755	6	AX451420 Sequence
26	107.8	3.6	2382	6	AX710668 Sequence
27	107.8	3.6	2388	6	AX710666 Sequence
28	107.8	3.6	2665	6	AX710664 Sequence
29	103.8	3.5	3450	3	AF145635 Drosophil
30	103.8	3.5	76299	2	AC019585 Drosophil
31	103.8	3.5	172363	3	AC009388 Drosophil
32	103.8	3.5	206190	3	AE003763 Drosophil
33	95.4	3.2	539	11	G98364
34	85.6	2.9	38145	3	U64847
35	75	2.5	40532	3	CBRG36D18
36	64.8	2.2	301995	1	AE016779 Pseudomon
37	61.4	2.0	10744	1	AE013659 Yersinia
38	61.4	2.0	214050	1	AJ414142 Yersinia
39	59.2	2.0	300029	1	AE012554 Xylella f
40	58	1.9	310029	1	AE016861 Pseudomon
41	56.6	1.9	299938	1	AE016800 Vibrio vu
42	56	1.9	11392	1	AE003940 Xylella f
43	55	1.8	10802	1	AE015425 Shigella
44	55	1.8	292497	1	AE016989 Shigella
45	54.6	1.8	221078	2	AC133100 Mus muscu

ALIGNMENTS

RESULT 1	AX420453	3003 bp	DNA	linear	PAT 18-JUN-2002
AX420453	Sequence 1 from Patent WO0216592.				
LOCUS	AX420453				
DEFINITION	AX420453				
ACCESSION	AX420453.1	GI:21524602			
VERSION	AX420453.1				
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1				
JOURNAL	Kapeller-Libermann, R. and Logan, T.J.				
	Atcr-1, a human acyltransferase and uses thereof				
	Patent: WO 0216592-A 1 28-FEB-2002;				

FEATURES		Millennium Pharmaceuticals, Inc. (US)
source	Location/Qualifiers	
	1..3003	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
CDS	341..2827	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD35486.1"	
	/db_xref="GI:21524603"	
	/translation="MDESALTLGTTIDVSLPHSSEYSGVRCKHTSEWEGCGRPRTVF RSATLWKESLMRKFVGRCCYSCTQSMDFKFNPSIPSLGURNVIYINETHRRH GWLARLSYVLFIEQWVHGKMFATNVTNLSRVQEAIAEVAABELNPDGSAQQQS KANVKKKAKKILQEMVATVSPAMIRLTGWLKLLFNPSFFWNIQIHKGQLEVMKAA ETNPLPLFVPHRSHDYLLLTFLFCHNIKAPYIASGNNLNIPIFSTLHKLGSPFI RRLEDPDGRKVDLYRALHGHIVELLRQOFLLEIFLEGTSRSGKTS CARAGLLSV VVDLSTNVIPDILLIPVGISYDRIIEGHNGEQLKPKKNESLWSVARGVIRMLRN YGCVRVDFAPRSLKEVLESQKPSALLSLEQALLPALPRSPDRQDEGRDTSIN ESRNATDESRLRILANLAHEILFTASKSCAIMSTHIVACLILYRHRQIDLSTLVED FYNKBEVLARDPLDFGNSGSDVVMHAIOLLGNCVTIHTSRNDEFFIPTSTVPSV PELNFVSNGLVDFIMEALIIACSLYAVLNKRGGLGPTSPNNLISQBLVRKAASLCY LLSNGSTISLPQCTFYQVCHETVKGFIQYGLITVAEHDQDIEDSPSLAEQQWKKLPE PLWSRDEDESDDFGEQRDCYLKVSQSKHQQFIFTLQRLGLPLEAYSSAIIKVVH NFGSPVPEPYLQKLHLITRNVNVAESATYCLVNQVAMFMDIGVFKETKOK RVSVLELSTFLPQCNKRLLEILISFVL"	
BASE COUNT	826 a 683 c 698 g 794 t	2 others
ORIGIN		
	Query Match 99.9%; Score 3001; DB 6; Length 3003;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 3003; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	TTGGGACACGAGCTGCTGGCGGGGACTCTTCTCAGGTTTACTGTGGAGCACCCAAAGTC 60
Db	1	TTGGGACACAGGCTGCTGGCGGGGACTCTTCTCAGGTTTACTGTGGAGCACCCAAAGTC 60
Qy	61	TGTCAGGCTCTGGCGGTGCAACAGGACCCAGAGGAACAGAGCTTGTCTTATTTCACCCA 120
Db	61	TGTCAGGCTCTGGCGGTGCAACAGGACCCAGAGGAACAGAGCTTGTCTTATTTCACCCA 120
Qy	121	CAGCTGGGACTGTCTTCTCCAGAGTCTCCATCAGCTTTTGTCTAATCGACTGATGGAAAT 180
Db	121	CAGCTGGGACTGTCTTCTCCAGAGTCTCCATCAGCTTTTGTCTAATCGACTGATGGAAAT 180
Qy	181	AATTCTCAACACCAACCAAGTCAAGGATACAGGACAGCGGCTCCCTGTGTATGGA 240
Db	181	AATTCTCAACACCAACCAAGTCAAGGATACAGGACAGCGGCTCCCTGTGTATGGA 240
Qy	241	CATTCTGACCCGAACTGATAGCTGAGTCTCTGAAGTTTATGTTATGAACAGAGAAC 300
Db	241	CATTCTGACCCGAACTGATAGCTGAGTCTCTGAAGTTTATGTTATGAACAGAGAAC 300
Qy	301	TTTTCATCCAGCAGCATGATTGGGAATTACACTTTGTGACATGGAATGATCTGCACGTGAC 360
Db	301	TTTTCATCCAGCAGCATGATTGGGAATTACACTTTGTGACATGGAATGATCTGCACGTGAC 360
Qy	361	CCTTGTGTACAATAGATGTTTCTTATCTGCCACATTCATCAGAATACAGTGTGTGTCGATG 420
Db	361	CCTTGTGTACAATAGATGTTTCTTATCTGCAATTCATCAGAATACAGTGTGTGTCGATG 420
Qy	421	TAAGCACACAAGTACAGGAATGGGTGAGTGTGGCTTTTAGACCCACCGTCTTCAGATCTGC 480
Db	421	TAAGCACACAAGTACAGGAATGGGTGAGTGTGGCTTTTAGACCCACCGTCTTCAGATCTGC 480
Qy	481	AACCTTTAAATGGAAGAAAGCCCTTAATGAGTCGGAAGGCCCAATTTGTTGGAAGATGTTG 540
Db	481	AACCTTTAAATGGAAGAAAGCCCTTAATGAGTCGGAAGGCCCAATTTGTTGGAAGATGTTG 540
Qy	541	TTACTCTCGACTCCCGCAGAGCTGGGACAAATTTTTTCAACCCCGATATCCCGTCTTTGGG 600
Db	541	TTACTCTCGACTCCCGCAGAGCTGGGACAAATTTTTTCAACCCCGATATCCCGTCTTTGGG 600

Db	2761	CATTTTCTACCTCAATGTCGAACCGACAAAACCTTCGTAGAAATATATTTCTGAGTTTGTGGT	2821
Qy	2821	GCTGTAGTAAACGCTGTGGCACTGCTGGCAAAATGAAGTCAATGAGATGAGTTCTCTGTGAGG	2880
Db	2821	GCTGTAGTAAACGCTGTGGCACTGCTGGCAAAATGAAGTCAATGAGATGAGTTCTCTGTGAGG	2880
Qy	2881	TACAGCTTCTGCTCTCAAGATTTGAAGTGCTTCGCGAGGGTCAAGCTCGCCCTGTGTC	2940
Db	2881	TACAGCTTCTGCTCTCAAGATTTGAAGTGCTTCGCGAGGGTCAAGCTCGCCCTGTGTC	2940
Qy	2941	CGAAGTGATCTCTGGAAGACAGTGCCTTCCTNCCTCCATGGATCTGAGATCTTCCGAGC	3000
Db	2941	CGAAGTGATCTCTGGAAGACAGTGCCTTCCTNCCTCCATGGATCTGAGATCTTCCGAGC	3000
Qy	3001	TTT 3003	
Db	3001	TTT 3003	
RESULT 2			
LOCUS	AX451419	3210 bp	DNA linear PAT 03-JUL-2002
DEFINITION	Sequence 26 from Patent WO0226950.		
ACCESSION	AX451419		
VERSION	AX451419.1	GI:21698428	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	1		
	Lal,P.G., Tang,Y.T., Yue,H., Burford,N., Gandhi,A.R., Warren,B.A.,		
	Yao,M.G., Tribouley,C.M., Baughn,M.R., Lee,E.A., Hafalia,A.J.,		
	Lu,Y., Griffin,J.A., Sanjanwala,M.S. and Ding,L.		
TITLE	Transferrases		
JOURNAL	Patent: WO 0226950-A 26 04-APR-2002;		
FEATURES	Incyte Genomics, Inc. (US)		
source	Location/Qualifiers		
	1..3210		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/notes="Incyte ID No: 1517230CB1"		
BASE COUNT	896 a 721 c 735 g 858 t		
ORIGIN			
Query Match 98.5%; Score 2958; DB 6; Length 3210;			
Best Local Similarity 99.7%; Pred.No.0;			
Matches 2983; Conservative 0; Mismatches 7; Indels 2; Gaps 2;			
Qy	12	GCTGTCGCGGGGGACTCTTCTGAGTTACTGTGGAGCACCACCAAGTCTGTGAGCTCT	71
Db	9	GCTGTCGCGGGGGACTCTTCTGAGTTACTGTGGAGCACCACCAAGTCTGTGAGCTCT	68
Qy	72	GGCGGTGCAAAACAGGCACCCAGAGAACAGACTTGTCTTATTCACCCACAGCTGGAC	131
Db	69	GGCGGTGCAAAACAGGCACCCAGAGAACAGACTTGTCTTATTCACCCACAGCTGGAC	128
Qy	132	TGCTTTCTCAGAGTCTCCATCAGCTTGTCTAATCGACTGATTTGGAATAATTTCTCAAA	191
Db	129	TGCTTTCTCAGAGTCTCCATCAGCTTGTCTAATCGACTGATTTGGAATAATTTCTCAAA	188
Qy	192	CACCAACAGTCAAGATACAGCAGCGGCTCCCTGTTGTATGGACATTTCTGCAC	251
Db	189	CACCAACAGTCAAGATACAGCAGCGGCTCCCTGTTGTATGGACATTTCTGCAC	248
Qy	252	CGAAATCTGATGCTGAGTCTCTGAAGTTTATGTTATGAAACAGAAGAACTTTTCATCCAG	311
Db	249	CGAAATCTGATGCTGAGTCTCTGAAGTTTATGTTATGAAACAGAAGAACTTTTCATCCAG	308
Qy	312	CACATGATTTGGAAATTACATTTGTGACATGGAATCTGCACGACCTTTGGTACAA	371
Db	309	CACATGATTTGGAAATTACATTTGTGACATGGAATCTGCACGACCTTTGGTACAA	368

QY	372	TAGATGTTTCTTATCTGCCACATTCATCAGATACAGTGTGGTTCGATGTAAAGCACACAA	431
Db	369	TAGATGTTTCTTATCTGCCACATTCATCAGATACAGTGTGGTTCGATGTAAAGCACACAA	428
QY	432	GTGAGGAATGGGTGAGTGTGGCTTTAGACCCACCGCTTTTCAGATCTGCACCTTTAAAT	491
Db	429	GTGAGGAATGGGTGAGTGTGGCTTTAGACCCACCATCTTCAGATCTGCACCTTTAAAT	488
QY	492	GGAAAGAAAGCCTTAATGAGTCGGAAGGCCATTTGTTGGAAGTGTGTACTCTCTGCA	551
Db	489	GGAAAGAAAGCCTTAATGAGTCGGAAGGCCATTTGTTGGAAGTGTGTACTCTCTGCA	548
QY	552	CTCCCGACAGCTGGGCAAAATTTTCAACCCCGAGTATCCCGTCTTTGGGTTTCGGAATG	611
Db	549	CTCCCGACAGCTGGGCAAAATTTTCAACCCCGAGTATCCCGTCTTTGGGTTTCGGAATG	608
QY	612	TTATTATATCAATGAAATTCACACAGACACCGCGGATGGCTTGCAAGACGCTTTCTTT	671
Db	609	TTATTATATCAATGAAATTCACACAGACACCGCGGATGGCTTGCAAGACGCTTTCTTT	668
QY	672	ACGTTCTTTTATCAAGAGGAGATGTGCATTAAGGCGATGTTTGGCCACCAATGTCACTG	731
Db	669	ACGTTCTTTTATCAAGAGGAGATGTGCATTAAGGCGATGTTTGGCCACCAATGTCACTG	728
QY	732	AAATGTGCTGAACAGCAGTAGAGTACAAGAGGCAATTCGAGAAGTGGCTGCTGAATTA	791
Db	729	AAATGTGCTGAACAGCAGTAGAGTACAAGAGGCAATTCGAGAAGTGGCTGCTGAATTA	788
QY	792	ACCCTGATGGTTCTGCCAGCAGCAATCAAAAGCCGTTAAACAAAGTGAAGAAAGACTA	851
Db	789	ACCCTGATGGTTCTGCCAGCAGCAATCAAAAGCCGTTAAACAAAGTGAAGAAAGACTA	848
QY	852	AAAGGATCTTCAAGAAATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGCGGT	911
Db	849	AAAGGATCTTCAAGAAATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGCGGT	908
QY	912	GGGTGCTGCTAAAATGTTTCAACAGCTTCTTTTGGAACTTCAAAATTCACAAAGGTCAAC	971
Db	909	GGGTGCTGCTAAAATGTTTCAACAGCTTCTTTTGGAACTTCAAAATTCACAAAGGTCAAC	968
QY	972	TTGAGATGGTTAAAGCTGCAACTGAGACGAATTTGCGGCTTCTGTTTCTACCAAGTTTCA	1031
Db	969	TTGAGATGGTTAAAGCTGCAACTGAGACGAATTTGCGGCTTCTGTTTCTACCAAGTTTCA	1028
QY	1032	GATCCCATATTTGACTATCTGCTGCTCATTTTCAATCTCTTCTGCCATAACATCAAGCAC	1091
Db	1029	GATCCCATATTTGACTATCTGCTGCTCATTTTCAATCTCTTCTGCCATAACATCAAGCAC	1088
QY	1092	CATACATGCTTCCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTGATCCATAAGC	1151
Db	1089	CATACATGCTTCCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTGATCCATAAGC	1148
QY	1152	TTGGGGCTTCTTTCATACGACGAAGGCTCGATGAACACACAGATGGACGGAAGATGTTTC	1211
Db	1149	TTGGGGCTTCTTTCATACGACGAAGGCTCGATGAACACACAGATGGACGGAAGATGTTTC	1208
QY	1212	TCATAGAGCTTTGCTCCATGGGCATATAGTTGAAATTAATTCGACAGCAGCAATTTCTGG	1271
Db	1209	TCATAGAGCTTTGCTCCATGGGCATATAGTTGAAATTAATTCGACAGCAGCAATTTCTGG	1268
QY	1272	AGATCTTCTGGAAGGCACACGTTCTAGGAGTGGAAACCTCTTGTGCTCGGGCAGGAC	1331
Db	1269	AGATCTTCTGGAAGGCACACGTTCTAGGAGTGGAAACCTCTTGTGCTCGGGCAGGAC	1328
QY	1332	TTTTGTGAGTGTGGTAGACTCTGTCTACCAATGTATCCAGACATCTTGATTAATAC	1391
Db	1329	TTTTGTGAGTGTGGTAGACTCTGTCTACCAATGTATCCAGACATCTTGATTAATAC	1388
QY	1392	CTGTTGGAATCTCCTATGATCGCAATTCGAAAGGTCACTCAATGTGTGAACAACTGGGCA	1451
Db	1389	CTGTTGGAATCTCCTATGATCGCAATTCGAAAGGTCACTCAATGTGTGAACAACTGGGCA	1448

QY	1452	AACCTAAGAAAGATGAGAGCCTGTGAGGTGTAGCAAGAGGTGTATTAGAAATGTTACGAA	1511
Db	1449	AACCTAAGAAAGATGAGAGCCTGTGAGGTGTAGCAAGAGGTGTATTAGAAATGTTACGAA	1508
QY	1512	AAAACTATGTTGTGTCCGAGTGGATTTTGCACAGCCATTTTCTTTAAAGGAATATTAG	1571
Db	1509	AAAACTATGTTGTGTCCGAGTGGATTTTGCACAGCCATTTTCTTTAAAGGAATATTAG	1568
QY	1572	AAAGCCAAAGTCAGAAAACGGGTGCTCTACTTTCTCCCTGGAGCAAGCGTGTCTTACCAG	1631
Db	1569	AAAGCCAAAGTCAGAAAACGGGTGCTCTACTTTCTCCCTGGAGCAAGCGTGTCTTACCAG	1628
QY	1632	CTATATCTTCTTTCAAGACCACAGTGTCTGTGATGAAGGTAGAGACACGTCCATTAATG	1691
Db	1629	CTATATCTTCTTTCAAGACCACAGTGTCTGTGATGAAGGTAGAGACACGTCCATTAATG	1688
QY	1692	AGTCCAGAAATGCAACAGATGAATCCCTACGAAGGAGGTGATTGCAAAATCTGGCTGAGC	1751
Db	1689	AGTCCAGAAATGCAACAGATGAATCCCTACGAAGGAGGTGATTGCAAAATCTGGCTGAGC	1748
QY	1752	ATATTCTTACTGCTAGCAAGTCTGCTGCCATTATGTCCACACATTTGGCTTGCC	1811
Db	1749	ATATTCTTACTGCTAGCAAGTCTGCTGCCATTATGTCCACACATTTGGCTTGCC	1808
QY	1812	TGCTCCTCTACAGACACAGCGAGGGAATTTGATCTCTCCACATTTGGTTCGAAGACTTCTTTG	1871
Db	1809	TGCTCCTCTACAGACACAGCGAGGGAATTTGATCTCTCCACATTTGGTTCGAAGACTTCTTTG	1868
QY	1872	TGATGAAAGAGGAAGTCTTGGCTCGTGATTTTGAACCTGGGGTCTCAGGAAATTCAGAAG	1931
Db	1869	TGATGAAAGAGGAAGTCTTGGCTCGTGATTTTGAACCTGGGGTCTCAGGAAATTCAGAAG	1928
QY	1932	ATGTAGTAATGCATGCCATACAGCTGCTGGGAAATTTGTCTCAATTCACCCACACTAGCA	1991
Db	1929	ATGTAGTAATGCATGCCATACAGCTGCTGGGAAATTTGTCTCAATTCACCCACACTAGCA	1988
QY	1992	GGAAACGATGAGTTTTTATCACCCCCAGACAACTGTCCCATCAGTCTTTCGAACTCAACT	2051
Db	1989	GGAAACGATGAGTTTTTATCACCCCCAGACAACTGTCCCATCAGTCTTTCGAACTCAACT	2048
QY	2052	TTTACAGCAATGGGGTACTTTCATGTCTTTATCATGAGGCGCATCATGCTTGAGGCTTT	2111
Db	2049	TTTACAGCAATGGGGTACTTTCATGTCTTTATCATGAGGCGCATCATGCTTGAGGCTTT	2108
QY	2112	ATGCGATTTCTGAACAGAGGGGACTGGGGGTCCCACTAGCACCCCACTTAACCTGATCA	2171
Db	2109	ATGCGATTTCTGAACAGAGGGGACTGGGGGTCCCACTAGCACCCCACTTAACCTGATCA	2168
QY	2172	GCCAGGAGCAGCTGGTGGGAAAGGGCCAGGCTGTGTACCTTCTCTCCAATGAAGGCA	2231
Db	2169	GCCAGGAGCAGCTGGTGGGAAAGGGCCAGGCTGTGTACCTTCTCTCCAATGAAGGCA	2228
QY	2232	CCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAACAGTAGGAAAGTTTA	2291
Db	2229	CCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAACAGTAGGAAAGTTTA	2288
QY	2292	TCCAGTATGGCATTTTACAGTGGCAGACGATGACAGGAGATATCAGTCTTAGTC	2351
Db	2289	TCCAGTATGGCATTTTACAGTGGCAGACGATGACAGGAGATATCAGTCTTAGTC	2348
QY	2352	TTGCTGAGCAGCAGTGGGCAAGAGACTTCCAGAACCTTTGTCTTGGAGAAAGTGAAG	2411
Db	2349	TTGCTGAGCAGCAGTGGGCAAGAGACTTCTTGAACCTTTGTCTTGGAGAAAGTGAAG	2408
QY	2412	AAGATGAACAGTGTGCTTTGGGAGGAACACGCGAGATTTGCTACCTGAAAGGTGAGCAAT	2471
Db	2409	AAGATGAACAGTGTGCTTTGGGAGGAACACGCGAGATTTGCTACCTGAAAGGTGAGCAAT	2468
QY	2472	CCAGGAGCAGCAGGATTTTATCACTTCTTACAGACTCTCTGGGCTTTGCTGGAG	2531
Db	2469	CCAGGAGCAGCAGGATTTTATCACTTCTTACAGACTCTCTGGGCTTTGCTGGAG	2528
QY	2532	CCTACAGCTCTGCTGCCATCTTTTGTTCACAACTTTCAGTGGTCTCTTCCAGAACCTGAGT	2591



```
Db 2529 CCTACAGCTCTGCTGCCATCTTTGTTCAACATTCAGTGGCTGCTTCCAGAACCTGAGT 2588
Qy 2592 ATCTGCAAAAGTTGCACAAATACCTTAATAACAGAACAGAAAGAAATGTCAGTATATG 2651
Db 2589 ATCTGCAAAAGTTGCACAAATACCTTAATAACAGAACAGAAAGAAATGTCAGTATATG 2648
Qy 2652 CTGAGAGTGCACATATTTGCTTTGGAAGATGCTGTGAAAATGTTTAAAGATATGGG 2711
Db 2649 CTGAGAGTGCACATATTTGCTTTGGAAGATGCTGTGAAAATGTTTAAAGATATGGG 2708
Qy 2712 TTTTCAAGGACCAACAAAGAGAGTCTGTTTGAAGTCTGTTTGAAGTCTGAGCAGCATTCTAC 2771
Db 2709 TTTTCAAGGACCAACAAAGAGAGTCTGTTTGAAGTCTGTTTGAAGTCTGAGCAGCATTCTAC 2768
Qy 2772 CTCAATGCAACCGACAAAATCTTAGAATATATTCAGTATTTGCTGCTAGGTAA 2831
Db 2769 CTCAATGCAACCGACAAAATCTTAGAATATATTCAGTATTTGCTGCTAGGTAA 2828
Qy 2832 CGTGTGGCAGTCTGTGGCAAAATGAAGTCTAGATGAGTCTCTTGTAGGTACCACTTCT 2891
Db 2829 CGTGTGGCAGTCTGTGGCAAAATGAAGTCTAGATGAGTCTCTTGTAGGTACCACTTCT 2888
Qy 2892 GGCCTCAAGAGTTGAAGTCTGCTTCCGAGGCTGAGGCTGCTGCTGCTGCTGCTGCT 2951
Db 2889 GGCCTCAAGAG-TTGAAGTCTGCTTCCGAGGCTGAGGCTGCTGCTGCTGCTGCTGCT 2946
Qy 2952 CCTGGAACAGAGTCTGCTTCCGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 3003
Db 2947 CCTGGAACAGAGTCTGCTTCCGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2998
```

```
RESULT 3
HSM804404 4918 bp mRNA linear PRI 30-APR-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp451P0819 (from clone DKFZp451P0819);
DEFINITION complete cds.
ACCESSION AL833093
VERSION AL833093.2 GI:30268382
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4918)
Bloecker H., Boecker M., Brandt P., Mewes H.W., Weil B., Amid C.,
Fobo G., Han M., Osanger A. and Wiemann S.
Direct Submission
Submitted (30-APR-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuerberg, GERMANY
On Apr 30, 2003 this sequence version replaced gi:21733684.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp451P0819) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1. 4918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="10q25.2"
/clone="DKFZp451P0819"
/tissue_type="cDNA-collection"
/clone_lib="451 (synonym: hlccl). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stages="adult"
1. 4918
gene
```

```
/gene="DKFZp451P0819"
200..2686
/gene="DKFZp451P0819"
/notes="glycerol-3-phosphate acyltransferase"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD89932.1"
/db_xref="GI:30268383"
/translations="MDESALTGLTIDVSLPHSEYSGVRCKTSEBEGECGRPTTIF
RSATLWKESIMSRKPFVGRCCYCTPQSDKFFNPSPISLGLRNVIYNETHRHR
GWLARRLSYVLFQERDVHKMFATNTVENLNSSRVOEBAIAEVAELNPDGSAQQS
KANKVAKAKRILQENWATVSPAMIKLTGVLKLFNFSFFNFQIHKGLQEMWKAAT
ETNLPFLFVHRSHDYLLELTILFCHNIKAPIYASGNLNPFIPTLHKLGCFI
RRRLDTPDKDVLVRAHLLGHIVELLRQQQFLEIFEGTRSRSGTSCARAGLLSV
VDTLTNTVPIIDILIPVGISYDRIIEGHVNGQLGPKKNESLWSVARGVIRMLKVN
YGCVRVDFAPFSLKEYLESOSKPVSALESLEQALLPAILPSRPPSAADGGRDTSN
ESRNATEDSARRRLIANLAHEHILFTASKSCAIMSTHIVACLLLYRHRQGLDLSLVED
FFVWKEVLARDLGLGNSDVMHAIQLGNCVTITHTSRNDEFTIDTSTPVS
FELNFSNGVLHVIMEAIIACSLYAVLNKRGGLGPTSTPNLISQOLVRKAASLCY
LLNBSGTISLPQCTFYQVCHETVGFIRYIGLTVAEHDDQEDISPSLAEQWDAKLPE
PLWSRDEDEDSDFGEORDCYLKVSQSEHOQIFLQRLGLPLEAYSSAALFVH
NFGSPVPEPEYLOKLHKYLTTRERNVAVAESATYCLVKNVKNFKDYGIVFKETQK
RVSYLESLSTFPLQCNRQKLELYLSFVVL"
polyA_site 4893
BASE COUNT 1379 a 1005 c 1075 g 1459 t
ORIGIN
Query Match 91.8%; Score 2755.4; DB 9; Length 4918;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2781; Conservative 0; Mismatches 8; Indels 2; Gaps 2;
Qy 213 GGCAGCAGCGCTCCCTGTTGATGACATCTTGCACCCGAAACTGATAGTCAGTCCT 272
Db 72 GGCAGCAGCGCTCCCTGTTGATGACATCTTGCACCCGAAACTGATAGTCAGTCCT 131
Qy 273 GAAGTTTTATGTTATGAAACAGAGAACTTTTCATCCCGACACATGATTTGGGAATTACAC 332
Db 132 GAAGTTTTATGTTATGAAACAGAGAACTTTTCATCCCGACACATGATTTGGGAATTACAC 191
Qy 333 TTTGTGACATGGATGAATCTGCACTGACCCCTGGTGTACATAGATGTTTCTTATCTGCCAC 392
Db 192 TTTGTGACATGGATGAATCTGCACTGACCCCTGGTGTACATAGATGTTTCTTATCTGCCAC 251
Qy 393 ATTCAATCAGAAATACAGTGTGTTGTCGATGTAAGACACACAAAGTAGGAATGGGTGAGTGTG 452
Db 252 ATTCAATCAGAAATACAGTGTGTTGTCGATGTAAGACACACAAAGTAGGAATGGGTGAGTGTG 311
Qy 453 GCTTTAGACCCACCGCTCTTCAGATCTGCAACTTTTAAATGAAAGAAAGCCTATGAGTC 512
Db 312 GCTTTAGACCCACCATCTTCAGATCTGCAACTTTTAAATGAAAGAAAGCCTATGAGTC 371
Qy 513 GGAAGAGCCATTTGTTGGAAGATGTTGTTATCTGCACTCCCGACAGCTGGGACAAAT 572
Db 372 GGAAGAGCCATTTGTTGGAAGATGTTGTTATCTGCACTCCCGACAGCTGGGACAAAT 431
Qy 573 TTTTCAACCCAGTATCCCGCTTTGGTGTGGGATGTTTATATATCAATCAAACTC 632
Db 432 TTTTCAACCCAGTATCCCGCTTTGGGTTTGGGAATGTTTATATCAATCAAACTC 491
Qy 633 ACACAAGACACCGCGATGGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGAGC 692
Db 492 ACACAAGACACCGCGATGGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGAGC 551
Qy 693 GAGATGTCATAAGGGCATGTTTGGCCACCAATGTGACTGAAAAATGTCTGAACAGCAGTA 752
Db 552 GAGATGTCATAAGGGCATGTTTGGCCACCAATGTGACTGAAAAATGTCTGAACAGCAGTA 611
Qy 753 GAGTACAGAGGCAATGTCAGAGTGGCTCTGAAATTAACCCCTGATGGTCTCTCCCGAGC 812
Db 612 GAGTACAGAGGCAATGTCAGAGTGGCTCTGAAATTAACCCCTGATGGTCTCTCCCGAGC 671
Qy 813 AGCAATCAAAAGCCGTTTAAACAAAGTGAAGAAAGCTAAAGAGGATTTCTTCAAGAAATGG 872
```

Db	672	AGCAATCAAAAGCGGTAAACAAAGTGAAAAAGAAAGCTAAAGGAATCTTCAAGAAATGG	731
Qy	873	TTGCCACTGCTCACCGGCAATGATCAGACTGACTGGTGGTGCCTTAAACTGTTC	932
Db	732	TTGCCACTGCTCACCGGCAATGATCAGACTGACTGGTGGTGCCTTAAACTGTTC	791
Qy	933	ACAGCTCTCTTTGGAAACATTCAAATTCACAAAGTCAACTTTGAGATGGTTAAAGCTGCAA	992
Db	792	ACAGCTCTCTTTGGAACTTTCAAATTCACAAAGTCAACTTTGAGATGGTTAAAGCTGCAA	851
Qy	993	CTGAGACGAATTTGCCGCTCTCTGTTCTACCAAGTTCATAGATCCCATATTTGACTATCTGC	1052
Db	852	CTGAGACGAATTTGCCGCTCTCTGTTCTACCAAGTTCATAGATCCCATATTTGACTATCTGC	911
Qy	1053	TGCTCACTTTCATCTCTCTGCGCATAAACATCAAGACCATACATTTGCTTACAGCAATA	1112
Db	912	TGCTCACTTTCATCTCTCTGCGCATAAACATCAAGACCATACATTTGCTTACAGCAATA	971
Qy	1113	ATCTCAACATCCCAATCTTCAGTACCTTTGATCCATAGCTTTGGGGCTTCTTCATACGAC	1172
Db	972	ATCTCAACATCCCAATCTTCAGTACCTTTGATCCATAGCTTTGGGGCTTCTTCATACGAC	1031
Qy	1173	GAAGGCTCGATGAACACCCAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATG	1232
Db	1032	GAAGGCTCGATGAACACCCAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATG	1091
Qy	1233	GGCATATAGTTGAATTTACTTCGACAGCAGCAATTTCTGGAGATCTTCTGGAAAGGCACAC	1292
Db	1092	GGCATATAGTTGAATTTACTTCGACAGCAGCAATTTCTGGAGATCTTCTGGAAAGGCACAC	1151
Qy	1293	GTTCTAGAGGTGGAAGAACTCTTGTGCTCGGCGAGGACTTTTGTGTCAGTTGTGCTAGATA	1352
Db	1152	GTTCTAGAGGTGGAAGAACTCTTGTGCTCGGCGAGGACTTTTGTGTCAGTTGTGCTAGATA	1211
Qy	1353	CTCTGTCTACCAATGTCCAGACATCTTGTATATACCTGTTGGAATCTCCTATGATC	1412
Db	1212	CTCTGTCTACCAATGTCCAGACATCTTGTATATACCTGTTGGAATCTCCTATGATC	1271
Qy	1413	GCATTTATCGAAGTCACTCAATGGTGAACAACTGGGCAACCTTAAGAAGATGAGAGCC	1472
Db	1272	GCATTTATCGAAGTCACTCAATGGTGAACAACTGGGCAACCTTAAGAAGATGAGAGCC	1331
Qy	1473	TGTGGAGTGTAGCAAGAGTGTATTAGATGTTTACGAAAACTATGTTGTGTCGAG	1532
Db	1332	TGTGGAGTGTAGCAAGAGTGTATTAGATGTTTACGAAAACTATGTTGTGTCGAG	1391
Qy	1533	TGGATTTTGCACAGCCATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGG	1592
Db	1392	TGGATTTTGCACAGCCATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGG	1451
Qy	1593	TGTCTGCTCTATCTTCTCGAGCAGCGTGTGTTACACAGCTATCTTCTTCAAGACCCA	1652
Db	1452	TGTCTGCTCTATCTTCTCGAGCAGCGTGTGTTACACAGCTATCTTCTTCAAGACCCA	1511
Qy	1653	GTGATGCTGCTGATGAGGTAGACACGTCCTTAATAGTCCAGAAATGCAACAGATG	1712
Db	1512	GTGATGCTGCTGATGAGGTAGACACGTCCTTAATAGTCCAGAAATGCAACAGATG	1571
Qy	1713	AATCCCTTACCAAGAGGTTGATTTGCAATCTGGCTGAGCATATTTCTATTCACCTGCTAGCA	1772
Db	1572	AATCCCTTACCAAGAGGTTGATTTGCAATCTGGCTGAGCATATTTCTATTCACCTGCTAGCA	1631
Qy	1773	AGTCTGTGCGCATTAATGTCACACACATTTGTGGTTCCTGCTCTCTACAGACAGGC	1832
Db	1632	AGTCTGTGCGCATTAATGTCACACACATTTGTGGTTCCTGCTCTCTACAGACAGGC	1691
Qy	1833	AGGGAATTTGATCTCTCCACATTTGTCGAAGACTTTCTTTGTGATGAAGAGGAAAGTCTG	1892
Db	1692	AGGGAATTTGATCTCTCCACATTTGTCGAAGACTTTCTTTGTGATGAAGAGGAAAGTCTG	1751
Qy	1893	CTCGTGATTTTGACCTGGGGTCTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATAC	1952

Db	1752	CTCGTGATTTTGACCTGGGGTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATAC	1811
Qy	1953	AGCTGCTGGGAAATTTGTGTCACAATCACCCACACTAGCAGGAAACGATGAGTTTTTATCA	2012
Db	1812	AGCTGCTGGGAAATTTGTGTCACAATCACCCACACTAGCAGGAAACGATGAGTTTTTATCA	1871
Qy	2013	CCCCCAGCAACACTGTGCCATCAGTCTTCGAACTCAACTTCTACAGCAATGGGTACTTTC	2072
Db	1872	CCCCCAGCAACACTGTGCCATCAGTCTTCGAACTCAACTTCTACAGCAATGGGTACTTTC	1931
Qy	2073	ATGCTCTTATCATGAGGCGCATCATAGCTTTCAGGCTTTATGCAATTCGAAACAGAGGG	2132
Db	1932	ATGCTCTTATCATGAGGCGCATCATAGCTTTCAGGCTTTATGCAATTCGAAACAGAGGG	1991
Qy	2133	GACTGGGGGTCCTACTAGCACCCACCTTAACTGATCAGCCAGGAGAGCTGTGCGGA	2192
Db	1992	GACTGGGGGTCCTACTAGCACCCACCTTAACTGATCAGCCAGGAGAGCTGTGCGGA	2051
Qy	2193	AGGCGGCCAGGCTGTGCTACTTCTCTCAATGAAGGCAACCATCTCAGCTTGCCTTGCAGA	2252
Db	2052	AGGCGGCCAGGCTGTGCTACTTCTCTCAATGAAGGCAACCATCTCAGCTTGCCTTGCAGA	2111
Qy	2253	CATTTTACCAAGTGTGCCATGAACACAGTAGGAAAGTTTATCCAGTATGCGCTTCTTACAG	2312
Db	2112	CATTTTACCAAGTGTGCCATGAACACAGTAGGAAAGTTTATCCGCTTATGCGCTTCTTACAG	2171
Qy	2313	TGCGAGCAGCAGTACAGCAGGAGATATCAGTCTCTAGTCTCTGCTGAGCAGAGTGGGACA	2372
Db	2172	TGCGAGCAGCAGTACAGCAGGAGATATCAGTCTCTAGTCTCTGCTGAGCAGAGTGGGACA	2231
Qy	2373	AGAAGCTTCCAGAACTTGTCTGTGGAGAAAGTGTATGAAGAGATGAAGACAGTACTTTG	2432
Db	2232	AGAAGCTTCCAGAACTTGTCTGTGGAGAAAGTGTATGAAGAGATGAAGACAGTACTTTG	2291
Qy	2433	GGGAGAAACAGCGAGATGCTACTCTGAGAGTGAAGCAATCCAAAGGAGCAGCAGAGTTTA	2492
Db	2292	GGGAGAAACAGCGAGATGCTACTCTGAGAGTGAAGCAATCCAAAGGAGCAGCAGAGTTTA	2351
Qy	2493	TCACCTTCTTACAGAGACTCTTGGGCTTCTGCTGAGGCTTACAGCTCTGCTGCCATCT	2552
Db	2352	TCACCTTCTTACAGAGACTCTTGGGCTTCTGCTGAGGCTTACAGCTCTGCTGCCATCT	2411
Qy	2553	TTGTTTCAACAATCTCAGTGGTCTCTTCCAGAACTCTGAGTATCTGCAAAAGTTGCAAAAT	2612
Db	2412	TTGTTTCAACAATCTCAGTGGTCTCTTCCAGAACTCTGAGTATCTGCAAAAGTTGCAAAAT	2471
Qy	2613	ACCTAATAACAGAAACAGAAAGATGTTGCAAGTATATGCTGAGAGTGCACATATTTGC	2672
Db	2472	ACCTAATAACAGAAACAGAAAGATGTTGCAAGTATATGCTGAGAGTGCACATATTTGC	2531
Qy	2673	TTGTGAGAAATGCTGTGAAATGTTTAAAGTATTTGGGGTTTTCAGGAGAGCCAAACAA	2732
Db	2532	TTGTGAGAAATGCTGTGAAATGTTTAAAGTATTTGGGGTTTTCAGGAGAGCCAAACAA	2591
Qy	2733	AGAGAGTGTCTGTTTGTAGAACTGAGCAGCCTTTTCTACCTCAATGCAACCCAGCAAAAC	2792
Db	2592	AGAGAGTGTCTGTTTGTAGAACTGAGCAGCCTTTTCTACCTCAATGCAACCCAGCAAAAC	2651
Qy	2793	TTCTAGAAATATTTCTGAGTTTGTGGTCTGTAGGTAAACGTTGGGCACTGTGTGGCAAT	2852
Db	2652	TTCTAGAAATATTTCTGAGTTTGTGGTCTGTAGGTAAACGTTGGGCACTGTGTGGCAAT	2711
Qy	2853	GAAGTTCATGAGTGTGCTTCTGAGTACAGGCTTCTGCTTCTGAGGTTTGAAGGTGC	2912
Db	2712	GAAGTTCATGAGTGTGCTTCTGAGTACAGGCTTCTGCTTCTGAGGTTTGAAGGTGC	2770
Qy	2913	CTTTCAGAGGGGTGAGGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT	2972
Db	2771	CATGCA - GGGTCAGGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT	2829
Qy	2973	NCCTCATCGATCTGAGATCTTCCAGCTTT 3003	
Db	2830	CCCTCCATGGATCTGTGATCTTCCAGCTCT 2860	

RESULT 4  
BC030783  
LOCUS  
DEFINITION Homo sapiens, KIAA1560 protein, clone MGC:26846 IMAGE:4821342,  
mRNA, complete cds.  
ACCESSION BC030783  
VERSION BC030783.1 GI:21315057  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Strausberg, R.  
Direct Submission  
Submitted (31-MAY-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapsb-remail.nih.gov](mailto:cgapsb-remail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
contact: [amandawebster@biology.org](mailto:amandawebster@biology.org)  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha  
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAP Plate: 33 Row: b Column: 19  
This clone was selected for full length sequencing because it  
passed the following selection criteria: GenomesScan gene  
prediction, similarity but not identity to protein.  
FEATURES  
source  
location  
1..3490  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="LocusID:57678"  
/db\_xref="taxon:9606"  
/clone="MGC:26846 IMAGE:4821342"  
/cissue\_type="Brain, hippocampus"  
/clone\_lib="NIH MGC\_95"  
/lab\_host="DH10B"  
/note="Vector: pBluescript"  
179..2665  
/codon\_start=1  
/product="KIAA1560 protein"  
/protein\_id="AAH30783.1"  
/db\_xref="GI:21315058"  
/translation="MDEVALTLGIDVSLPHSSEYSGVRCKHTSEWGEGRPTIP  
RSATLKWKEISLMSRKEPVGRCYCTPQSDWKPEPTSIPLSLGRNVIYNEHTRR  
GWLARRLSYLFIERDVKHGMFPAINTVENSRSQEAIAEVAELNPDGSAQQS  
KAVNKKAKRILOEMVATSPAMIRLTGWLLKLFNSFPWNIQIHKGQLEVMKAAAT  
ETNPLFLPVHRSHIDYLLTLFLFCHNIKAPYIASGNLNLIPIFTLHLKLGFFI  
RRLDDEYDGRKDLVYRALLHGHIVELLRQQOFLIELEGRNRSKGTSCARAGLLSV  
VVDLTSTNVI PDILLIPVGISYDRI IEHNGEOLGPKKNESLWSVARGWMLRKN  
YGCVRVPAQPSKLEVESQKPSVALLSLEQALLPALIPSPSDAEDGRDTSIN  
ESRNATESLRRRLIANLAELHIFASKMSITHIVACLLLYRHQGDLSLTVSD  
FVWKVEEDLDFDGLSGNSDVMHAIQLLGNCTVITHTSRNDEFFITPSTVPVS  
FELNPSYNGVILFIMEAIIACSLYAVLNKRGLGPTSTPENLISQBLVRKAALCY  
LLSNGEITSLPCQTFYOVCHETGVGFIOYGLITVAEHDDQEDISPLABQOWDKLPE  
PLWSRDEEDSDSDFGEORDCYLVKVSQKHOQFIFLQRLGLGLEAYSSAIFVH  
NFSGPVPEPEYLQKLHKYLIITRERNVANVAESATYCLVRKNVQKFDIGVFKETKQ  
RVSVELSSTFLPQCNRQKLLLEYILISFVVL"  
1002 a 734 c 790 g 964 t  
BASE COUNT  
ORIGIN

Query Match 91.6%; Score 2752.2; DB 9; Length 3490;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 2779; Conservative 0; Mismatches 10; Indels 2; Gaps 2;  
QY 213 GGCAGCAGCGCTCCCTCTGTTGTATGGACATTTCTGCAACCCGAAACTGATAGCTAGTCT 272  
DB 51 GGCAGCAGCGCTCCCTCTGTTGTATGGACATTTCTGCAACCCGAAACTGATAGCTAGTCT 110  
QY 273 GAAGTTTTATGTTATGAAACAGAGAACTTTTCATCCCGACGACATGATTGTTGGAAATTACAC 332  
DB 111 GAAGTTTTATGTTATGAAACAGAGAACTTTTCATCCCGACGACATGATTGTTGGAAATTACAC 170  
QY 333 TTTGTGACATGGATGAATCTGCACTGACCCCTTGGTCAACAATAGATGTTTCTTATCTGCCAC 392  
DB 171 TTTGTGACATGGATGAATCTGCACTGACCCCTTGGTCAACAATAGATGTTTCTTATCTGCCAC 230  
QY 393 ATTCAATCAGATACAGTGTGGTGGTGAAGACACACAGTGAAGAAATGGGGTGAAGTGTG 452  
DB 231 ATTCAATCAGATACAGTGTGGTGGTGAAGACACACAGTGAAGAAATGGGGTGAAGTGTG 290  
QY 453 GCTTTAGACCCACCGCTCTTCAGATCTCAACTTTTAAATGGAAGAAAGCCCTAATAGATC 512  
DB 291 GCTTTAGACCCACCGCTCTTCAGATCTCAACTTTTAAATGGAAGAAAGCCCTAATAGATC 350  
QY 513 GGAAGAGCCCATTTGTTGGAGATGTTGTTACTCTCTGCACCTCCCAGAGCTGGGACAAAT 572  
DB 351 GGAAGAGCCCATTTGTTGGAGATGTTGTTACTCTCTGCACCTCCCAGAGCTGGGACAAAT 410  
QY 573 TTTTCAACCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAATC 632  
DB 411 TTTTCAACCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAATC 470  
QY 633 ACACAGACACCGCGATGCTTCAGAGCGCTTTCAGAGCGCTTCTAGTCTTTTATTTCAAGAGC 692  
DB 471 ACACAGACACCGCGATGCTTCAGAGCGCTTTCAGAGCGCTTCTAGTCTTTTATTTCAAGAGC 530  
QY 693 GAGATGTGCATAGGGCATGTTTGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTA 752  
DB 531 GAGATGTGCATAGGGCATGTTTGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTA 590  
QY 753 GAGTACAAGAGCGCAATTCGAGAAGTGGCTGCTGAATTTAAACCCCTGATGGTTCGCCAGC 812  
DB 591 GAGTACAAGAGCGCAATTCGAGAAGTGGCTGCTGAATTTAAACCCCTGATGGTTCGCCAGC 650  
QY 813 AGCAATCAAAAGCGGTTTAAAGAGTGAAGAAAGCTTAAAGGATTTCTTCAAGAAATGG 872  
DB 651 AGCAATCAAAAGCGGTTTAAAGAGTGAAGAAAGCTTAAAGGATTTCTTCAAGAAATGG 710  
QY 873 TTGCCACTGTCTCACCGGCAATGATCAGACTGCTGGGTGGGTGCTGCTAAACCTGTTCA 932  
DB 711 TTGCCACTGTCTCACCGGCAATGATCAGACTGCTGGGTGGGTGCTGCTAAACCTGTTCA 770  
QY 933 ACAGCTTTCTTTTGGAAATTTCAATTCACAAAGTCAACTTTGAGATGGTGTAAAGCTGCAA 992  
DB 771 ACAGCTTTCTTTTGGAAATTTCAATTCACAAAGTCAACTTTGAGATGGTGTAAAGCTGCAA 830  
QY 993 CTGAGACGAATTTGCCCGCTTCTGTTTCTACAGTTTCATAGATCCCATATTGACTATCTGC 1052  
DB 831 CTGAGACGAATTTGCCCGCTTCTGTTTCTACAGTTTCATAGATCCCATATTGACTATCTGC 890  
QY 1053 TGCTCACTTTTCATTTCTTCTGCTCCATAAATCAAGACCAATACATTCATTCAGGCAATA 1112  
DB 891 TGCTCACTTTTCATTTCTTCTGCTCCATAAATCAAGACCAATACATTCATTCAGGCAATA 950  
QY 1113 ATCTCAACATCCCAATCTTCAGTACCTTGTATCCATAAGCTTTGGGGGCTTCTTCAACGAC 1172  
DB 951 ATCTCAACATCCCAATCTTCAGTACCTTGTATCCATAAGCTTTGGGGGCTTCTTCAACGAC 1010  
QY 1173 GAAGGCTCGATGAACACACAGATGGAGGAAAGATGTTCTCTATAGAGCTTTGCTCCATG 1232  
DB 1011 GAAGGCTCGATGAACACACAGATGGAGGAAAGATGTTCTCTATAGAGCTTTGCTCCATG 1070

QY 1233 GGCATATAGTGAATTAATTCGACAGCAGCAATTCCTTGAGATCTTCTCGAAGGCACAC 1292  
DB 1071 GGCATATAGTGAATTAATTCGACAGCAGCAATTCCTTGAGATCTTCTCGAAGGCACAC 1130  
QY 1293 GTTCTAGGAGTGGAATAAATCTCTTGCTCGGAGAGCAATTTTGTGTCAGATTTGTGGTAGATA 1352  
DB 1131 GTTCTAGGAGTGGAATAAATCTCTTGCTCGGAGAGCAATTTTGTGTCAGATTTGTGGTAGATA 1190  
QY 1353 CTCTGTCTACCAATGTCCATCCAGACATCTTGATATATACCTGTTGGAATCTCTCTATGATC 1412  
DB 1191 CTCTGTCTACCAATGTCCATCCAGACATCTTGATATATACCTGTTGGAATCTCTCTATGATC 1250  
QY 1413 GCATTATCGAAGTCACTACAATGGTGAAACAACTGGGCAAACTTAAGAAATAGAGAGCC 1472  
DB 1251 GCATTATCGAAGTCACTACAATGGTGAAACAACTGGGCAAACTTAAGAAATAGAGAGCC 1310  
QY 1473 TGTGGAGTGTAGCAAGAGTGTATATAGAAATGTTCGAAAAAATCTATGTTGTGTCGGAG 1532  
DB 1311 TGTGGAGTGTAGCAAGAGTGTATATAGAAATGTTCGAAAAAATCTATGTTGTGTCGGAG 1370  
QY 1533 TGGATTTTGCACAGCCATTTTCTTAAAGCAATATTTAGAAAGCCAAAGTCAGAAACCCG 1592  
DB 1371 TGGATTTTGCACAGCCATTTTCTTAAAGCAATATTTAGAAAGCCAAAGTCAGAAACCCG 1430  
QY 1593 TGTCTGTCTACTTTTCCCTGGAGCAAGCTGTTTACCCAGCTATATCTTCTCAAGACCCA 1652  
DB 1431 TGTCTGTCTACTTTTCCCTGGAGCAAGCTGTTTACCCAGCTATATCTTCTCAAGACCCA 1490  
QY 1653 GTGATGCTGTGATGAAGGTAGAGACACGTCCTCAATTAATAGTCCAGAAATGCACAGATG 1712  
DB 1491 GTGATGCTGTGATGAAGGTAGAGACACGTCCTCAATTAATAGTCCAGAAATGCACAGATG 1550  
QY 1713 AATCCCTACGAAGAGGTGAATTCGAAATCTGCTGAGCATATCTATTCACTGCTAGCA 1772  
DB 1551 AATCCCTACGAAGAGGTGAATTCGAAATCTGCTGAGCATATCTATTCACTGCTAGCA 1610  
QY 1773 AGTCCTGTGCAATATGTCACACACATGTCGCTGCTGCTCTCTACAGACACAGGC 1832  
DB 1611 AGTCCTGTGCAATATGTCACACACATGTCGCTGCTGCTCTCTCTACAGACACAGGC 1670  
QY 1833 AGGGAATGTATCTCTCCATGTCGAAGACTTCTTTGTGATGAAGAGAGAGTCTCTGG 1892  
DB 1671 AGGGAATGTATCTCTCCATGTCGAAGACTTCTTTGTGATGAAGAGAGAGTCTCTGG 1730  
QY 1893 CTGCTGATTTTGACCTGGGGTCTCAGGAATTCAGNAGATAGTAATGTCATGCCATAC 1952  
DB 1731 CTGCTGATTTTGACCTGGGGTCTCAGGAATTCAGNAGATAGTAATGTCATGCCATAC 1790  
QY 1953 AGCTGTGGGAAATTTGTGCACAAATCACCCACACTAGCAGGAACGATGATTTTATCA 2012  
DB 1791 AGCTGTGGGAAATTTGTGCACAAATCACCCACACTAGCAGGAACGATGATTTTATCA 1850  
QY 2013 CCCCCAGCAACTGTGCCCATCAGTCTTGAATCTCAACTTTCTACAGCAATGGGGTACTTC 2072  
DB 1851 CCCCCAGCAACTGTGCCCATCAGTCTTGAATCTCAACTTTCTACAGCAATGGGGTACTTC 1910  
QY 2073 ATGCTTTTATCATGGAGCCATCATAGCTTGCAGCCCTTTATGCAATGCTCAACAGAGG 2132  
DB 1911 ATGCTTTTATCATGGAGCCATCATAGCTTGCAGCCCTTTATGCAATGCTCAACAGAGG 1970  
QY 2133 GACTGGGGGTCCCATCTAGCACCCACCTAACTGTATCAGCCAGGAGCAGCTGGTGGGA 2192  
DB 1971 GACTGGGGGTCCCATCTAGCACCCACCTAACTGTATCAGCCAGGAGCAGCTGGTGGGA 2030  
QY 2193 AGCGGGCCAGCTGTGCTACTCTTCTTCCAAATGAAGSCACCATCTCACTGCCCTGGCAGA 2252  
DB 2031 AGCGGGCCAGCTGTGCTACTCTTCTTCCAAATGAAGSCACCATCTCACTGCCCTGGCAGA 2090  
QY 2253 CATTTTACCAAGTCTGCCATGAACAGTAGAAGTTTATCCAGTATGGAATCTTACAG 2312  
DB 2091 CATTTTACCAAGTCTGCCATGAACAGTAGAAGTTTATCCAGTATGGAATCTTACAG 2150  
QY 2313 TGGCAGAGCAGATGACCGAAGAGATATCAGTCTCTAGTCTTGTGCTGAGCAGCAGTGGGACA 2372

DB 2151 TGGCAGACCATGATGACCGAAGATATCAGTCTTAGTCTTGTGAGCAGCAGTGGGACA 2210  
QY 2373 AGAAGCTTCCAGAACCTTTTGTCTTGGAGAGTGAAGAGATGAAGACAGTACTTTG 2432  
DB 2211 AGAAGCTTCCAGAACCTTTTGTCTTGGAGAGTGAAGAGATGAAGACAGTACTTTG 2270  
QY 2433 GGGAGAAACAGCGAGATTCCTACCTGAAGGTGAGCCAAATCCCAAGGAGCAGCAGCTTTA 2492  
DB 2271 GGGAGAAACAGCGAGATTCCTACCTGAAGGTGAGCCAAATCCCAAGGAGCAGCAGCTTTA 2330  
QY 2493 TCACCTTTCTTACAGAGACTCTTGGGCCCTTGTCTGGAGGCCCTACAGCTCTGTGCGCATCT 2552  
DB 2331 TCACCTTTCTTACAGAGACTCTTGGGCCCTTGTCTGGAGGCCCTACAGCTCTGTGCGCATCT 2390  
QY 2553 TTGTTTCAACAATTCAGTGTCTCTTCCAGAACCTGAGTATCTGCAAAAAGTTGCACAAAT 2612  
DB 2391 TTGTTTCAACAATTCAGTGTCTCTTCCAGAACCTGAGTATCTGCAAAAAGTTGCACAAAT 2450  
QY 2613 ACCTAATAACCAAGAAAGAAATGTTGTCAGTATATGCTGAGAGTGCACATATTGTC 2672  
DB 2451 ACCTAATAACCAAGAAAGAAATGTTGTCAGTATATGCTGAGAGTGCACATATTGTC 2510  
QY 2673 TTGTGAAGAAATCTGTGAAATTTTAAAGATATTGGGGTTTTCAGAGGAGACCAACAAA 2732  
DB 2511 TTGTGAAGAAATCTGTGAAATTTTAAAGATATTGGGGTTTTCAGAGGAGACCAACAAA 2570  
QY 2733 AGAGAGTGTCTGTTTGTAGAACTGAGCAGACATTTTCTACCTCAATGCAACCGACAAAAC 2792  
DB 2571 AGAGAGTGTCTGTTTGTAGAACTGAGCAGACATTTTCTACCTCAATGCAACCGACAAAAC 2630  
QY 2793 TTCTAGAAATATTTCTGAGTGTGTTGTTGCTGTAGTAACTGTTGCTGAGTGCACAAAT 2852  
DB 2631 TTCTAGAAATATTTCTGAGTGTGTTGTTGCTGTAGTAACTGTTGCTGAGTGCACAAAT 2690  
QY 2853 GAAGGTGATGAGATGATTTCTTGTAGTGTACAGCTTCTGGCTCAAGAGTTGAAGGTGC 2912  
DB 2691 GAAGGTGATGAGATGATTTCTTGTAGTGTACAGCTTCTGGCTCAAGAG-TTGAAGGTGC 2749  
QY 2913 CTTTCGAGGGGTGAGCGCTGCGCTGTCGAGTGTATCTCTGGAAGACAGTGCCTTCT 2972  
DB 2750 CATCGCA-GGGTCAGCGCTGCGCTGTCGGAAGTGTATCTCTGGAAGACAGTGCCTTCT 2808  
QY 2973 NCCTCCATCGATCTGAGATCTTCCAGCTTT 3003  
DB 2809 CCATCCATGATCTGTGATCTTCCAGCTCT 2839

## RESULT 5

HSM804372

LOCUS

DEFINITION

complete cds.

AL833061

VERSION

AL833061.2

GI:30268346

Homo sapiens (human)

Homo sapiens

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

HSM804372 6387 bp mRNA linear PRI 30-APR-2003  
Homo sapiens mRNA; cDNA DKFP451B115 (from clone DKFP451B115);  
complete cds.

AL833061  
AL833061.2 GI:30268346

Homo sapiens (human)

Homo sapiens

ORGANISM

REFERENCE

1 (bases 1 to 6387)

Bloeker, H., Boecher, M., Brandt, P., Mewes, H.W., Weill, B., Amid, C.,  
Fobo, G., Han, M., Osanger, A. and Wiemann, S.

Direct Submission

Submitted (30-APR-2003) MIPS, Ingolstaedter Landstr.1, D-85764

Neuherberg, GERMANY

On Apr 30, 2003 this sequence version replaced gi:21733651.

Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKPZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by GSF (National Research Centre for Biotechnology Ltd.,

Braunschweig/Germany) within the cDNA sequencing consortium of the

German Genome Project.

This clone (DKFP451B115) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

```
FEATURES
    source            1..6387
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /map="10q25.2"
                        /clone="DKFp451B1115"
                        /tissue_type="cDNA-collection"
                        /clone_lib="451 (synonym: hlccl). Vector pSport1; host
DH10B; sites NotI + SalI"
                        /dev_stage="adult"
    gene              1..6387
                        /gene="DKFp451B1115"
    CDS                197..2683
                        /gene="DKFp451B1115"
                        /note="glycerol-3-phosphate acyltransferase"
                        /codon_start=1
                        /product="hypothetical protein"
                        /protein_id="CAD89967.1"
                        /db_xref="GI:30268347"
                        /translation="WDSSALTLGIDVYLSHSSEYSGVRCKHTSEWEGCGFRPTVP
RSATLWKESLMRSRRPVGRCYSCSTPQSDWKFPNPSIPSLGRNVIYINETHR
GWLARRLSYVLFIOERDVHKGFMATNTENVLSRVQEAIAEVAEILNPDGSAQOQS
KAVNKVKKAKRIQEMVATVSPAMRLTGMVLLKLFNSFWNIQIHKGQLEWVKAAT
ETNLLPLFVHRSHIDYLLATLFIKFNKIPVIAAGNNLNIPIFSTLHLKGGFT
RRLEDETPDKGRKDVLYRALLGHVILLRQOQFLEIFLETRSRSKTSRCARGLLSV
VDTLNVFIDILLIPVIGISYRIEGHYNGEQKPKNESLWSVARGVIRMLRKN
YGCVRVDPAPQFSLKEVLSQSPKVSALLSLEQALLPAILPSRPSDAADGRDTSIN
ESRNATDESLEARRDIANLAHEILFTASKSCAIMSTHIVACLLLYRHRQGDLSLVED
PFVKEEVLARDFDLQSGNEDVVMHAIQLLGNCTVITHTSRNDEFFIPTSTVPSV
PELNFYNGVLHVRITMEAIACSLYAVLNKEGLGPTSPNLLISQQLVRKAASLCY
LLSNEATISLSCQFYQVCHETVKFYQYGLITVAEHDDQEDISPLAEQWDRKKLPE
LLSWRSEDSDFGEEQRDCYLVKSQSEKHQFIFLORLLGLPLLEAYSSAAIFVH
NFSQVEPEPEYLQKLHKYLIITRERNVAVYAESATYCLVKNVAMFKDIGVFKETQK
RVSVLELSSTPLPQCNROKLLEYLSFVVL"
    polyA_signal       6346..6351
                        /gene="DKFp451B1115"
    polyA_site         6364
                        /gene="DKFp451B1115"
BASE COUNT          1865 a 1239 c 1381 g 1902 t
ORIGIN
Query Match          91.6%; Score 2752.2; DB 9; Length 6387;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2779; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
Qy 213 GGCAGCAGCGCTCCCTGTTGTATGGACATTCGACCCGAACTGATAGCTGAGTCCT 272
Db      |||
Qy 69  GGCAGCAGCGCTCCCTGTTGTATGGACATTCGACCCGAACTGATAGCTGAGTCCT 128
Db      |||
Qy 273 GAACTTTTATGTTATGAACAGAGAAGCTTTTCATCCAGCAGCATGATTGGGAATTCAC 332
Db      |||
Qy 129 GAACTTTTATGTTATGAACAGAGAAGCTTTTCATCCAGCAGCATGATTGGGAATTCAC 188
Qy 333 TTTGTGACATGATGAATCTGCATGACCCCTTGGTACAAATAGATGTTCTTATCTGCCAC 392
Db      |||
Qy 189 TTTGTGACATGATGAATCTGCATGACCCCTTGGTACAAATAGATGTTCTTATCTGCCAC 248
Qy 393 ATTCAATCAGAAATACAGTGTGGTTCATGATGAAGCACAAGTGAGGAATGGGGTTCAGTGTG 452
Db      |||
Qy 249 ATTCAATCAGAAATACAGTGTGGTTCATGATGAAGCACAAGTGAGGAATGGGGTTCAGTGTG 308
Qy 453 GCTTTAGACCCACCGCTCTTCAGATCTGCAACTTTAAATAAGGAAGAAAGCCCTAATGAGTC 512
Db      |||
Qy 309 GCTTTAGACCCACCGCTCTTCAGATCTGCAACTTTAAATAAGGAAGAAAGCCCTAATGAGTC 368
Qy 513 GGAAGAGCCCAATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAAT 572
Db      |||
Qy 369 GGAAGAGCCCAATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAAT 428
Db      |||
```

1653	Qy	GTGATGCTGCTGATGAAGGTAGAGACAGCTTCATTATATGAGTCCAGAAATGCAACAGATG	1712
1509	Db	GTGATGCTGCTGATGAGGGTAGAGACAGCTCCATTATATGATGAGTCAGAAATGCAACAGATG	1568
1713	Qy	AATCCCTACGAAGGAGGTGATATGCAAAATCTGGCTGAGCATAATCTATTTCACTGCTAGCA	1772
1569	Db	AATCCCTACGAAGGAGGTGATATGCAAAATCTGGCTGAGCATAATCTATTTCACTGCTAGCA	1628
1773	Qy	AGTCCTGTGCCATTATGTCCACACACATTTGTGGCTTGCCTCTCTACAGACACAGGC	1832
1629	Db	AGTCCTGTGCCATTATGTCCACACACATTTGTGGCTTGCCTCTCTACAGACACAGGC	1688
1833	Qy	AGGAAATGTATCTCTCCATATGTFGTGGAAGACTTCTTTTGTGATGAAGAAGAGTCTCTGG	1892
1689	Db	AGGAAATGTATCTCTCCATATGTFGTGGAAGACTTCTTTTGTGATGAAGAAGAGTCTCTGG	1748
1893	Qy	CTCGTGATTTTGACCTGGGGTCTCAGGAATTTAGAGATGTAGTAAATGCATGCCATAC	1952
1749	Db	CTCGTGATTTTGACCTGGGGTCTCAGGAATTTAGAGATGTAGTAAATGCATGCCATAC	1808
1953	Qy	AGTGTCTGGGAAATTTGTGTCACAATCACCCACACTAGCAGGAAACGATGAGTTTTTATCA	2012
1809	Db	AGTGTCTGGGAAATTTGTGTCACAATCACCCACACTAGCAGGAAACGATGAGTTTTTATCA	1868
2013	Qy	CCCCAGCAACAATGTCCCACTAGTCTTGGAACCTCAACTTCTACAGCAATGGGGTACTTC	2072
1869	Db	CCCCAGCAACAATGTCCCACTAGTCTTGGAACCTCAACTTCTACAGCAATGGGGTACTTC	1928
2073	Qy	ATGTCITTTATCATGGAGGCCATCATAGCTTGCAGCCTTTATGCAGTTCTGAAACAAGAGGG	2132
1929	Db	ATGTCITTTATCATGGAGGCCATCATAGCTTGCAGCCTTTATGCAGTTCTGAAACAAGAGGG	1988
2133	Qy	GACTGGGGGTCCCACATAGCACCCCACTTAACCTGTATCAGCCAGGAGCAGCTGTGTGGGA	2192
1989	Db	GACTGGGGGTCCCACATAGCACCCCACTTAACCTGTATCAGCCAGGAGCAGCTGTGTGGGA	2048
2193	Qy	AGCGGCCAGCTGTGTACTCTTCTCCAAATGAAGCACCATCTCATGCTTGCACAG	2252
2049	Db	AGCGGCCAGCTGTGTACTCTTCTCCAAATGAAGCACCATCTCATGCTTGCACAG	2108
2253	Qy	CATTTTACCAAGTCTGCCATGAAACAGTAGAAGATTTATCCAGTATGGCAATCTTTACAG	2312
2109	Db	CATTTTACCAAGTCTGCCATGAAACAGTAGAAGATTTATCCAGTATGGCAATCTTTACAG	2168
2313	Qy	TGCAGAGCAGATGACAGGAGATATCAGTCTCTAGTCTTGTGTGACAGCAGTGGGACA	2372
2169	Db	TGCAGAGCAGATGACAGGAGATATCAGTCTCTAGTCTTGTGTGACAGCAGTGGGACA	2228
2373	Qy	AGAAGCTTCCAGAACCTTTGCTTGGAGAAGTGAATGAAGAGATGAAGACAGATGACTTTG	2432
2229	Db	AGAAGCTTCCAGAACCTTTGCTTGGAGAAGTGAATGAAGAGATGAAGACAGATGACTTTG	2288
2433	Qy	GGGAGGAACAGCGAGATTTGCTAAGAGGTGAGCCAAATCCAGGAGCAACAGCAGTTTA	2492
2289	Db	GGGAGGAACAGCGAGATTTGCTAAGAGGTGAGCCAAATCCAGGAGCAACAGCAGTTTA	2348
2493	Qy	TCACTTCTTACAGAGACTCCTTGGGCTTTGCTGGAGGCTACAGCTCTGCTGCCATCT	2552
2349	Db	TCACTTCTTACAGAGACTCCTTGGGCTTTGCTGGAGGCTACAGCTCTGCTGCCATCT	2408
2553	Qy	TTGTTCACAACTTCAGTGGTCTGTCCAGAACTGAGTATCTGCAAAAGTTGCACAAAT	2612
2409	Db	TTGTTCACAACTTCAGTGGTCTGTTCAGAACTGAGTATCTGCAAAAGTTGCACAAAT	2468
2613	Qy	ACCTAATAACAGAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCACATATTTGTC	2672
2469	Db	ACCTAATAACAGAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCACATATTTGTC	2528
2673	Qy	TTGTGAAGATGCTGTGAAAATGTTTAAAGATATTTGGGGTTTTCAAGGAGACCAACAAA	2732
2529	Db	TTGTGAAGATGCTGTGAAAATGTTTAAAGATATTTGGGGTTTTCAAGGAGACCAACAAA	2588
2733	Qy	AGAGAGTGTCTGTTTTTAAAGACTGAGCAGACTTTTCTACCTCAATGCAACCGACAAAAC	2792

Db	2589	AGAGAGTCTGTTTGTAGAACTGAGCAGCACTTTTCTACTCAATGCAACCGACAAAAC	2648
Qy	2793	TTCTAGAAATATATTCGTAGTGTGTTGGTGTAGGTAAAGCTGTGGCACTGCTGGCAAAAT	2852
Db	2649	TTCTAGAAATATATTCGTAGTGTGTTGGTGTAGGTAAAGCTGTGGCACTGCTGGCAAAAT	2708
Qy	2853	GAAGGTCATGAGATGAGTTCCTTGTAGTACCAAGCTTCTGGCTCAAGAGTTTGAAGGTGC	2912
Db	2709	GAAGGTCATGAGATGAGTTCCTTGTAGTACCAAGCTTCTGGCTCAAGAG-TTGAAGGTGC	2767
Qy	2913	CTTTCGAGGGGTGAGCCCTGCTCCCTGTCCTGGAAGTATCTCCTGGAAACAAAGTGCCTTCT	2972
Db	2768	CGTGCAC-GGGTCAGGCCCTCCCTGTCCCGAGGTGATCTCCTGGAAACAAAGTGCCTTCT	2826
Qy	2973	NCCTCCATGATCTGAGATCTTCCCAAGCTTT	3003
Db	2827	CCCTCCATGATCTGTGATCTTCCAGCTCT	2857
RESULT 6			
AX431391			
LOCUS	AX431391	2487 bp	DNA linear PAT 28-JUN-2002
DEFINITION	Sequence 24 from Patent WO0240666.		
ACCESSION	AX431391		
VERSION	AX431391.1	GI:21656238	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Enkins, D.K., Winther, M.D., Haardt, M., Goldberg, Y.P., Nwaka, S.O., Ponton, A., Allen, S.J., de Antueno, R.J. and Knickle, L.C.		
TITLE	Fat regulated genes, uses thereof, and compounds for modulating same		
JOURNAL	Patent: WO 0240666-A 24 23-MAY-2002;		
FEATURES	XENON GENETICS INC (CA)		
source	Location/Qualifiers		
	1..2487		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	706 a 548 c 570 g 663 t		
ORIGIN			
Query Match	82.8%;	Score 2485.4;	DB 6; Length 2487;
Best Local Similarity	100.0%;	Pred. NO. 0;	
Matches 2486;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	341	ATGGATGAATCTGCACCTGACCTTGGTGTAGTATGATGTTTCTTATCTGCCACATTCATCA	400
Db	1	ATGGATGAATCTGCACCTGACCTTGGTGTAGTATGATGTTTCTTATCTGCCACATTCATCA	60
Qy	401	GAATACAGTCTGGTTCGATTAAGCACACAAGTGAGGAATGGGGTGGGTGAGTGGCTTTAGA	460
Db	61	GAATACAGTCTGGTTCGATTAAGCACACAAGTGAGGAATGGGGTGGGTGAGTGGCTTTAGA	120
Qy	461	CCCAAGTCTTCAGATCTGCACTTTAAATGAAAGAAAGCCCTATAGTGGGAAAGG	520
Db	121	CCCAAGTCTTCAGATCTGCACTTTAAATGAAAGAAAGCCCTATAGTGGGAAAGG	180
Qy	521	CCATTTGTTGGAAGTGTGTTACTCTGACCTCCCGAGCTGGGCAAAATTTTCAAC	580
Db	181	CCATTTGTTGGAAGTGTGTTACTCTGACCTCCCGAGCTGGGCAAAATTTTCAAC	240
Qy	581	CCAGTATCCCGCTTTTGGGTTTGGGAATGTTATTTATATCAATGAACTCACACAAGA	640
Db	241	CCAGTATCCCGCTTTTGGGTTTGGGAATGTTATTTATATCAATGAACTCACACAAGA	300
Qy	641	CACCGGATGGCTTGCACAGCCTTTCTTAAGTCTTTTATTTCAAGAGCGAGATGTG	700
Db	301	CACCGGATGGCTTGCACAGCCTTTCTTAAGTCTTTTATTTCAAGAGCGAGATGTG	360







Db 1741 ATCATGAGGCCATCATAGCTTGGAGCCCTTTATGAGTCTTGAAACAAGAGGGGACTGGGG 1800  
Qy 2141 GGTCCCACTAGCACCCCACTAACTGATACGACGAGGAGAGCTGGTGGGAAGCGCGCC 2200  
Db 1801 GGTCCCACTAGCACCCCACTAACTGATACGACGAGGAGAGCTGGTGGGAAGCGCGCC 1860  
Qy 2201 AGCCTGTGCTACCTTCTCTCCAAATGAAGGACCATCTCAGTGGCTTGGCAGACATTTTAC 2260  
Db 1861 AGCCTGTGCTACCTTCTCTCCAAATGAAGGACCATCTCAGTGGCTTGGCAGACATTTTAC 1920  
Qy 2261 CAAAGTCTGCATGAACAGTAGTGAAGTTTATCCAGTATGGCATTTCTACAGTGGCAGAG 2320  
Db 1921 CAAAGTCTGCATGAACAGTAGTGAAGTTTATCCAGTATGGCATTTCTACAGTGGCAGAG 1980  
Qy 2321 CACGATGACGAGGAAGATATCAGTCTCTAGTCTTGTGTCAGCAGCAGTGGGACAAAGAGCTT 2380  
Db 1981 CACGATGACGAGGAAGATATCAGTCTCTAGTCTTGTGTCAGCAGCAGTGGGACAAAGAGCTT 2040  
Qy 2381 CCAGAACCTTTGCTCTTGAGAGAGTATGAAGAAGATGAAGACAGTGAATTTGGGGAGGAA 2440  
Db 2041 CCAGAACCTTTGCTCTTGAGAGAGTATGAAGAAGATGAAGACAGTGAATTTGGGGAGGAA 2100  
Qy 2441 CAGCGAGATTGCTACCTGAAGTGAGCCAAATCCAGAGGAGCACCAGCAGTTTATCACCTTC 2500  
Db 2101 CAGCGAGATTGCTACCTGAAGTGAGCCAAATCCAGAGGAGCACCAGCAGTTTATCACCTTC 2160  
Qy 2501 TTACAGAGACTCTTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2560  
Db 2161 TTACAGAGACTCTTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2220  
Qy 2561 AACTTCAGTGGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAAATACCTAATA 2620  
Db 2221 AACTTCAGTGGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAAATACCTAATA 2280  
Qy 2621 ACCGAAACAGAAAGAAATGTTGCAATATGCTGAGAGTGCACATATGTTCTGTGAAG 2680  
Db 2281 ACCGAAACAGAAAGAAATGTTGCAATATGCTGAGAGTGCACATATGTTCTGTGAAG 2340  
Qy 2681 AATGCTGTGAAAATCTTTAAGGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTG 2740  
Db 2341 AATGCTGTGAAAATCTTTAAGGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTG 2400  
Qy 2741 TCTGTTTATAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGCAAAAACTTCTAGAA 2800  
Db 2401 TCTGTTTATAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGCAAAAACTTCTAGAA 2460  
Qy 2801 TATATCTGAGTTTGTGGTCTG 2824  
Db 2461 TATATCTGAGTTTGTGGTCTG 2484

RESULT 8  
HSM803772 HSM803772 5689 bp mRNA linear PRI 13-MAY-2003  
LOCUS Homo sapiens mRNA; cDNA DKFp313H071 (from clone DKFp313H071).  
DEFINITION AL832464  
ACCESSION AL832464  
VERSION 1  
KEYWORDS GI:21733030  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 5689)  
AUTHORS Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,  
Wewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and  
Wiemann, S.  
Direct Submission  
Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANY  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by DKFZ (German Cancer Research Center,  
Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.  
This clone (DKFp313H071) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
information about the clone and the sequencing project is available  
at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

Location/Qualifiers  
1..5689  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFp313H071"  
/tissue\_type="human fetal skin"  
/clone\_lib="313 (synonym: hlcc2). Vector pTriplEx2; host  
DH10B; sites SfiIA + SfiIB"  
/dev\_stage="fetal"  
polyA\_site 5654  
BASE COUNT 1521 a 1149 c 1256 g 1763 t  
ORIGIN

Query Match 74.7%; Score 2241.8; DB 9; Length 5689;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2246; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Qy 211 CAGCAGCAGCGCTCCCTGTTGTATGGACATTTCTGCACCCGAACTGATAGCTGAGTC 270  
Db 37 CAGCAGCAGCGCTCCCTGTTGTATGGACATTTCTGCACCCGAACTGATAGCTGAGTC 96  
Qy 271 CTGAAGTTTATGTTATGAAACAGAGAATCTTTCATCCAGCACATGTTTGGGAATTAC 330  
Db 97 CTGAAGTTTATGTTATGAAACAGAGAATCTTTCATCCAGCACATGTTTGGGAATTAC 156  
Qy 331 ACTTTGTGACATGATGAATCTGCACCTGACCTTGGTACATAGATGTTTCTTATCTGCC 390  
Db 157 ACTTTGTGACATGATGAATCTGCACCTGACCTTGGTACATAGATGTTTCTTATCTGCC 216  
Qy 391 ACATTTCATCAATATACAGTGTGTCGATGTAAGCACACACAGTGAAGGAGTGGGTGAGTG 450  
Db 217 ACATTTCATCAATATACAGTGTGTCGATGTAAGCACACACAGTGAAGGAGTGGGTGAGTG 276  
Qy 451 TGGCTTTAGACCCACCGCTTTCAGATCTGAACTTTTAAATGGAAGAACCCCTAAATGAG 510  
Db 277 TGGCTTTAGACCCACCGCTTTCAGATCTGAACTTTTAAATGGAAGAACCCCTAAATGAG 336  
Qy 511 TCGAAAAAGGCCATTTGTTGAAGATGTTGTTACTCTGCACTCCCGAGCTCGGACAA 570  
Db 337 TCGAAAAAGGCCATTTGTTGAAGATGTTGTTACTCTGCACTCCCGAGCTCGGACAA 396  
Qy 571 ATTTTTCACCCCACTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAC 630  
Db 397 ATTTTTCACCCCACTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAC 456  
Qy 631 TCACACAGACACCGCGGATGGCTTTCAGAACGCTTCTTACGTTCTTTTATTTCAAGA 690  
Db 457 TCACACAGACACCGCGGATGGCTTTCAGAACGCTTCTTACGTTCTTTTATTTCAAGA 516  
Qy 691 GCGAGATGTCATAAAGGCGCATGTTTGCACCAATGTCACATAATGCTGGAACAGCAG 750  
Db 517 GCGAGATGTCATAAAGGCGCATGTTTGCACCAATGTCACATAATGCTGGAACAGCAG 576  
Qy 751 TAGAGTCAAGAGGCAATTCAGAAAGTGGCTGCTGGAATTTAAACCTGATGTTCTGCCCA 810  
Db 577 TAGAGTCAAGAGGCAATTCAGAAAGTGGCTGCTGGAATTTAAACCTGATGTTCTGCCCA 636  
Qy 811 GCAGCAATCAAAAGCGCTTAAACAAAGTGAAGAAAGAGCTTAAAGAGATTTCTCAAGAAAT 870  
Db 637 GCAGCAATCAAAAGCGCTTAAACAAAGTGAAGAAAGAGCTTAAAGAGATTTCTCAAGAAAT 696  
Qy 871 GGTTCGCACTGTCTCAGCGGCAATGATCAGACTGCTGGGTGGGTGCTGCTTAAACTGTT 930  
Db 697 GGTTCGCACTGTCTCAGCGGCAATGATCAGACTGCTGGGTGGGTGCTGCTTAAACTGTT 756  
Qy 931 CAACAGCTTCTTTTGGAAACATTCAAAATTCACAAAGGTCACCTTGAGATGGTTAAAGCTGC 990

757 CAACAGCTTCTTTTGGAAACATTCAAATTCACAAAGGTCACCTTGAGATGTTAAAGCTGC 816  
991 AACTGAGACGAATTTGCGGCTTCTGTTTCTACCAAGTTCATAGATCCCATATTCAGCTATCT 1050  
817 AACTGAGACGAATTTGCGGCTTCTGTTTCTACCAAGTTCATAGATCCCATATTCAGCTATCT 876  
1051 GCTGCTCACCTTTCATCTCTCTGCCATAACATCAAAAGCACCACATACATCTGCTTCAGGCAA 1110  
877 GCTGCTCACCTTTCATCTCTCTGCCATAACATCAAAAGCACCACATACATCTGCTTCAGGCAA 936  
1111 TAATCTCAACATCCCAATCTTCAGTACCTTTGATCCATAAAGCTTTGGGGGCTTCTTCATACG 1170  
937 TAATCTCAACATCCCAATCTTCAGTACCTTTGATCCATAAAGCTTTGGGGGCTTCTTCATACG 996  
1171 ACGAGGCTCGATGAACACACAGATGGAACGAAAGATGTTCTCTATAGAGCTTTGCTTCCA 1230  
997 ACGAGGCTCGATGAACACACAGATGGAACGAAAGATGTTCTCTATAGAGCTTTGCTTCCA 1056  
1231 TGGSCATATAGTTGAATTTACTTCGACAGCAGCAATCTTCGAGATCTTCTCGAAGGCAC 1290  
1057 TGGSCATATAGTTGAATTTACTTCGACAGCAGCAATCTTCGAGATCTTCTCGAAGGCAC 1116  
1291 ACGTTCTAGAGTGGAACCACTTCTGCTCGGCGCAGGACTTTTGTGTCAGTTGTGTTAGA 1350  
1117 ACGTTCTAGAGTGGAACCACTTCTGCTCGGCGCAGGACTTTTGTGTCAGTTGTGTTAGA 1176  
1351 TACTCTGCTACCAATGTCATCCAGACATCTTGTGATTAATACCTTTGGAATCTCCTATGA 1410  
1177 TACTCTGCTACCAATGTCATCCAGACATCTTGTGATTAATACCTTTGGAATCTCCTATGA 1236  
1411 TCCATTTATCGAAGTCACTACATGTTGGAACCACTTGGGCAACCTTAAGAGATGAGAG 1470  
1237 TCGCATTTATCGAAGTCACTACATGTTGGAACCACTTGGGCAACCTTAAGAGATGAGAG 1296  
1471 CCTGTGAGTGATGAGAGAGTGTTATTAGAATGTTACGAAACCACTTATGTTGTGTCG 1530  
1297 CCTGTGAGTGATGAGAGAGTGTTATTAGAATGTTACGAAACCACTTATGTTGTGTCG 1356  
1531 AGTGGATTTGACAGGCACTTTTCTTAAAGGAATATTTAGAAGCCAAAGTCAGAAAC 1590  
1357 AGTGGATTTGACAGGCACTTTTCTTAAAGGAATATTTAGAAGCCAAAGTCAGAAAC 1416  
1591 GGTGCTGCTCTACTTTTCCCTGAGCAAGGTTGTTACAGCTATATCTTCTTCAAGACC 1650  
1417 GGTGCTGCTCTACTTTTCCCTGAGCAAGGTTGTTACAGCTATATCTTCTTCAAGACC 1476  
1651 CAGTGATGCTGCTCATGAAGGTAGAGACACGTCCTTAATTAATGAGTCCAGAAATGCAACAG 1710  
1477 CAGTGATGCTGCTCATGAAGGTAGAGACACGTCCTTAATTAATGAGTCCAGAAATGCAACAG 1536  
1711 TGAATCCCTACGAAGGAGTTGATGCAATCTGGCTGAGCATATCTTATTCATCTGCTAG 1770  
1537 TGAATCCCTACGAAGGAGTTGATGCAATCTGGCTGAGCATATCTTATTCATCTGCTAG 1596  
1771 CAGTCTGCTGCTCATGAGGTTGCTGCAACATGTTGGCTTGGCTGCTCTCTACAGACACAG 1830  
1597 CAGTCTGCTGCTCATGAGGTTGCTGCAACATGTTGGCTTGGCTGCTCTCTACAGACACAG 1656  
1831 GCAGGGAATTTGATCTCTCCCATTTGTCGAAGACTTCTTTTGTGATGAAGAGGAAGTCTCT 1890  
1657 GCAGGGAATTTGATCTCTCCCATTTGTCGAAGACTTCTTTTGTGATGAAGAGGAAGTCTCT 1716  
1891 GGCTCGTGAATTTGACCTGGGTTCTCAGGAAATTCAGAGATGTTAGTAAATGATGCGCAT 1950  
1717 GGCTCGTGAATTTGACCTGGGTTCTCAGGAAATTCAGAGATGTTAGTAAATGATGCGCAT 1776  
1951 ACAGCTCTGGGAAATTTGTCGAATTCACCCACCTAGCAGGAAACGATCAGTCTTTTAT 2010  
1777 ACAGCTCTGGGAAATTTGTCGAATTCACCCACCTAGCAGGAAACGATCAGTCTTTTAT 1836  
2011 CACCCCGACGACAACTGCTCCCATCAGTCTTTCGAACCTCAACTTCTTACAGCAATGGGTACT 2070

Db 1837 CACCCCGACGACAACTGTCCCATCAGTCTTGAACCTCAACTTCTACAGCAATGGGTACT 1896  
Qy 2071 TCATGTCCTTTATCATCGAGGSCCATCATAGCTTGCAGCCCTTTATGCACTTCTGAACAAGAG 2130  
Db 1897 TCATGTCCTTTATCATCGAGGSCCATCATAGCTTGCAGCCCTTTATGCACTTCTGAACAAGAG 1956  
Qy 2131 GGGACTTGGGGGGTCCCACTAGCACCCCACTAACCTGATCAGCCAGGAGCAGCTGGTGG 2190  
Db 1957 GGGACTTGGGGGGTCCCACTAGCACCCCACTAACCTGATCAGCCAGGAGCAGCTGGTGG 2016  
Qy 2191 GAAGGGGCGAGCTGTGTGTACTCTTCTCCAAATGAAGGACCATCTCACTGCTTGGCCA 2250  
Db 2017 GAAGGGGCGAGCTGTGTGTACTCTTCTCCAAATGAAGGACCATCTCACTGCTTGGCCA 2076  
Qy 2251 GACATTTTACCAGTCTGCCATGAACAGTGAAGATTTATCCAGTATGGCATCTTCTAC 2310  
Db 2077 GATATTTTACCAGTCTGCCATGAACAGTGAAGATTTATCCAGTATGGCATCTTCTAC 2136  
Qy 2311 AGTGGCAGAGCAGCATGACGAGGAAGATATCAGTCTTCTAGTCTTGTGAGCAGCAGTGGGA 2370  
Db 2137 AGTGGCAGAGCAGCATGACGAGGAAGATATCAGTCTTCTAGTCTTGTGAGCAGCAGTGGGA 2196  
Qy 2371 CAAGAAGCTTCAGAACCTTTGTCTTGGAGAGTGTATGAAGAGATGAAGACAGTGACTT 2430  
Db 2197 CAAGAAGCTTCCTGAACCTTTGTCTTGGAGAGTGTATGAAGAGATGAAGACAGTGACTT 2256  
Qy 2431 TGGGGAGGACAGCGAGATGCTTACCTGAAGT 2463  
Db 2257 TGGGGAGGACAGCGAGATGCTTACCTGAAGT 2289

RESULT 9  
AB046780  
LOCUS Homo sapiens mRNA for KIAA1560 protein, partial cds.  
DEFINITION  
ACCESSION AB046780  
VERSION AB046780.1 GI:10047184  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Nagase, T., Kikuno, R., Nakayama, M., Hirose, M. and Ohara, O.  
TITLE Prediction of the coding sequences of unidentified human genes. XVIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro  
JOURNAL DNA Res. 7 (4), 273-281 (2000)  
MEDLINE 20450683  
PUBMED 10997877  
REFERENCE 2 (bases 1 to 5674)  
AUTHORS Ohara, O., Nagase, T. and Kikuno, R.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute, Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnaifn@kazusa.or.jp, URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913, Fax: 81-438-52-3914)

FEATURES  
source Location/Qualifiers  
1..5674  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="fh19959"  
/note="vector:pBluescriptII SK plus"  
1..5674  
/gene="KIAA1560"  
<1..1989  
/gene="KIAA1560"  
/notes="Start codon is not identified."  
/codon\_start=1  
/product="KIAA1560 protein"  
/protein\_id="BAB13386.1"











Db 2098 AAAACTACGGCTACGTCAGATGGATTTTGGACAGCCATTTTCTTGAAGGAATTTAG 2157  
Qy 1572 AAAGCAAGAGTCAGAAACCGGTGCTGCTCTACTTTCCCTGGAGCAACGCTTTGTTACCAG 1631  
Db 2158 AAGGCCAGATCAGAAACCTGTATCTGCCCTTTTCTCTGGAGCAACGCTTTTACCAG 2217  
Qy 1632 CTATACCTTCCCTCAAGACCCAGTGTGCTGATGAAGGTAGAGACACGCTCCATTAATG 1691  
Db 2218 CGATCTTCTTCAAGACCCGATGATGTTGCTGATGAACATCAAGACCTATCCAGTAACG 2277  
Qy 1692 AGTCAGAAATGCAACAGATGAATCCCTACGAAGGAGGTGATTCGAATCTGGCTGAGC 1751  
Db 2278 AGTCAGAAACCCAGCAGACGAGCTTCCGACGAGGCTGATTCGAACCTGGCTGAGC 2337  
Qy 1752 ATATTCTATTCACTGTAGCAAGTCTGTGCCATTATTCACACACACATTTGGCTTGCC 1811  
Db 2338 ACATTCTCTACGCCAGCAAGTCTGTGCTATCATGTCCACCCACATTTGGGCTGTC 2397  
Qy 1812 TGCTCTCTACAGACACAGGACGGAATGATCTCTCCACATTTGGTCCGAACCTTTTG 1871  
Db 2398 TGCTCTCTACAGACACAGGACGGAATCCATCTCTCCACGCTTTGGGAAGACTTTCTTG 2457  
Qy 1872 TGATGAAGAGAGTCTGCTGCTGATTTTGAACCTGGGTCTCAGGAATTCAGAG 1931  
Db 2458 TGATGAAGAGAGTCTGCTGCTGATTTTGAACCTGGGTCTCAGGAATTCAGAG 2517  
Qy 1932 ATGTAGTAATGCATGCCATACAGCTGTGGGAAATTTGTGCACATCACCCACACTAGCA 1991  
Db 2518 ATGTCTCATGATGCTATTACGCTTCTGGGAACTGTGTACATCACCCACAGGACA 2577  
Qy 1992 GGAACGATGATTTTATACCCGCCAGCAAACTGTCCCACTAGTCTTCCAACTCAACT 2051  
Db 2578 GGAAGATGATTTTATTAATCTCCAGCAAACTGTCCCGTCACTTTTGAACCTCAACT 2637  
Qy 2052 TCTACAGCAATGGGTACTTCTATGCTTTATCATGAGGACCATCATAGCTTGGACCTTT 2111  
Db 2638 TCTACAGCAATGGGTACTTCTATGCTTTATCATGAGGACCATCATAGCTTGGACCTTT 2697  
Qy 2112 ATGCAGTCTTGAACAAAGAGGAGTGGGGGTCCCACTAGCACCCACCACTAACCTGATCA 2171  
Db 2698 ATGCAGTCTTGAATGAAGGTGCTCTGAGAGGTCCGTGGAGGCTCGGCAACCTGATCA 2757  
Qy 2172 GCCAGGACGCTGTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2231  
Db 2758 GCCAGGACGCTGTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2817  
Qy 2232 CCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCAAGTCTGCAAGTCTGCAAGTCT 2291  
Db 2818 CCATCTCTGCTTGCAGACATTTTACCAAGTCTGCAAGTCTGCAAGTCTGCAAGTCT 2877  
Qy 2292 TCCAGTATGGCATTTTACAGTGGCAGACGATGACGAGGAGATATCAGTCTAGTC 2351  
Db 2878 TCCAGTATGGCATTTTACAGTGGCAGACGATGACGAGGAGATATCAGTCTAGTC 2937  
Qy 2352 TTGCTGACGACGATGGGACAGAGCTTCCAGAACCTTTGCTTGGAGAGTATGATGAG 2411  
Db 2938 TTGAGAACACGACGATGGGACAGAGCTTCCGAGAC---TGAACCTGGAGAGTACAGG 2994  
Qy 2412 AAGATGAAGACGATGATTTGGGAGGAGAACAGCGAGATTTGCTAGTGAAGTACGCAAT 2471  
Db 2995 AAGATGAAGACGATGATTTGGGAGGAGAACAGCGAGATTTGCTAGTGAAGTACGCAAT 3054  
Qy 2472 CCAAGGAGCAGCAGGATTTATCACTTTTACAGAGACTCTTGGGCTTTGGGAGG 2531  
Db 3055 CCAAGGAGCAGCAGGATTTATCACTTTTACAGAGACTCTTGGGCTTTGGGAGG 3114  
Qy 2532 CCTACAGCTCTGCTGCCATCTTTGTTTCAAACTTCAAGTCTGCTGTTCCAGAACCTGAGT 2591  
Db 3115 CCTACAGCTCTGCTGCCATCTTTGTTTCAAACTTCAAGTCTGCTGTTCCAGAACCTGAGT 3174  
Qy 2592 ATCTGCAAAAGTTGCAAAATACCTTATTAACAGAACAGAAAGAAATTTGTCAGATATG 2651  
Db 3175 ACCTGCAAAAGTTGCAAAATACCTTATTAACAGAACAGAAAGAAATTTGTCAGATATG 3234

Qy 2652 CTGAGAGTGCACATATTTCTTGTGAAGATGCTGTGAATGTTTAAAGGATATTTGGG 2711  
Db 3235 CTGAGAGTGCACATATTTCTTGTGAAGATGCTGTGAATGTTTAAAGGATATTTGGG 3294  
Qy 2712 TTTTCAAGGAGACCAACAAAGAGAGTGTCTGTTTGTAGAACTGAGCAGCACTTTTCTAC 2771  
Db 3295 TTTTCAAGGAGACCAACAAAGAGAGTGTCTGTTTGTAGAACTGAGCAGCACTTTTCTAC 3354  
Qy 2772 CTCAATGCAACCGACAAACCTCTAGAAATATTTCTGAGTTTGTGGTGTAGGTAA 2831  
Db 3355 CTCAATGCAACCGACAAACCTCTAGAAATATTTCTGAGTTTGTGGTGTAGGTAA 3414  
Qy 2832 CGTGTGCACTGTCTGGCAAAATGAAAGTCTAGAGATGATTTCTTGTAGTACCAAGCTTCT 2891  
Db 3415 CTCCT--CAGCACTATCTCGAGCGAGAGCTGATGATGCTTTCGACCTCCAGTCTGA 3472  
Qy 2892 GGCTCAAGAGTTTGAAGTGTCTTCGCAAGGCTCAGGCTGCTGCTGCTGCTGCTGCT 2951  
Db 3473 CCCAGA---GTGCAAGTATCTTGGCATGCCCCA-GCCTGCTGCTGCTGCTGCTGCT 3527  
Qy 2952 CCTGGAAGACAGTGCCTTCTNCTCCATGATCTGAGATCTTCCGAGC 3000  
Db 3528 CCAGAAAGACAGTGCCTTCTGCCCCGTGAGCCTGCTGCTGCTGCTGCTGCTGCTGCT 3576

## RESULT 12

LOCUS AX401972 2646 bp DNA linear PAT 06-JUN-2002  
DEFINITION Sequence 1648 from Patent W00210453.  
ACCESSION AX401972  
VERSION AX401972.1 GI:21338152

KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus

REFERENCE 1 Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and Elashoff, M.R.  
AUTHORS Molecular toxicology modeling  
TITLE Patent: WO 0210453-A 1648 07-FEB-2002;  
JOURNAL Gene Logic, Inc. (US)  
FEATURES Location/Qualifiers

1..2646  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/note="EMBL/GenBank Accession No. NM\_017274"  
BASE COUNT 670 a 689 c 670 g 617 t  
ORIGIN

Query Match 63.1%; Score 1896.2; DB 6; Length 2646;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2126; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 319 TTTGGGAATTACACTTTTGTGACATGATGATCTGCACTGACCTTGGTCAATAGATGT 378  
Db 138 TTTGGGACTTCCAGTTTGTGATGATGAGGAGTCTTCTAGTGACAAATTGGCAATAGACGT 197  
Qy 379 TTCTTATCTGCCAATTCATCAGAAATACAGTTTGGTTCGATGTAAGCACACAAGTGAAGA 438  
Db 198 TTCTTATCTGCCAATTCATCAGAAATACAGTTTGGGCGATGTAACACAGAAATGAGA 257  
Qy 439 ATGGGTGAGTGTGGCTTTTAGACCCACCTGCTTTCAGATCTGCAACTTTAAATGGAAGA 498  
Db 258 CTGGTTGATGTTGGCTTTCAACCTTCTTTCAGATCCGACGCTGAAATGGAAGA 317  
Qy 499 AAGCTTAATGATCGGAAAGGCAATTTGTTGGAAGATGTTTACTTCTGCACTCCCA 558  
Db 318 GAGCTCATGAGCGGAAAGGCGCTTCTGTTGGGAAAGGTTGCTTATTCATGACGCTCA 377

Qy	559	GAGCTGGGACAAATTTTTTCAACCCAGTAGTATCCCGCTCTTTGGGTTTGGGAATGTTATTTA	618
Db	378	GAGCTGGGAAGGTTTTTCAACCCAGTAGTATCCCACTCTCGGTTTGGGAATGTTATTTA	437
Qy	619	TATCAATGAACACTCACACAGACACACCGGATGGCTTTCGAAGACGCCCTTTCTTACGTTCT	678
Db	438	TATCAATGAAACTCACACAAGGCACCCAGGATGGCTGGCAAGCGGCTTTCTTACATCCT	497
Qy	679	TTTTTATTTCAAGAGCGAGATGTGCAATAAGGGCATGTTTGGCCACCAATGTGACTGAAATGT	738
Db	498	TTTTTGTTCAGAGGCGCATGTCTCACAGGGCATGTTTGGCCACCAGTATCACTGACATGT	557
Qy	739	GCTGAACAGCAGTAGTAGTACAAGAGGCAATTGCAAGAGTGGCTGCTGCAATTAACCCCTGA	798
Db	558	ACTGAATAGCAGCAGAGTCCAAGAGGCAATTGCTGAAGTGGCTGCAGAAATTGAACCCGGA	617
Qy	799	TGGTTCTGCCAGCAGCAATCAAAGCGGTTAACAAAGTGAAGAAAGAAAGCTAAAGGAT	858
Db	618	TGGATCTGCCAGCAGCAGTCCAAGGCATCCAGAAAGTGAAGAAAGAAAGCCAGGAAGAT	677
Qy	859	TCTTCAAGAAATGGTTGCCACTGTCTCACCGGCAATCATCAGACTGACTGGGTGGGTCT	918
Db	678	CCTCCAGAAATGGTTGCTACAGTCTCCCGGGATGATCAGGCTGACTGGCTGGGTGTT	737
Qy	919	GCTAAACTGTTCACAGCTCTTTTTTGGAACTTCAAATTCACAAAGGTCAACTTGAGAT	978
Db	738	ACTAAAGCTCTTCAACAGCTTCTTCTGGAACATTCAGATTCACAAGGTCACACTTGAGAT	797
Qy	979	GGTTAAAGCTGCAACTGAGAGCAATTTGCGGCTTCTGTTCTACCAAGTTCAATAGATCCCA	1038
Db	798	GGTGAAGCTGCAACTGAGAGCAATCTGCGGCTCTTGTCTGCGGTGCACAGATCCCA	857
Qy	1039	TATTGACTATCTGCTGCTCACTTTCTTCTCTGCAATAACATCAAGCACCACATACAT	1098
Db	858	CATCGACTACTGCTGCTCACTTCTCTCTGCAATAACATCAAGCTCCATACAT	917
Qy	1099	TGCTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTGATCCATAGCTTGGGG	1158
Db	918	CGCCTCGGCAACAACCTCAACATCCCACTCTTCAGTACCTTGATTCACAAGCTTGGGG	977
Qy	1159	CTTCTTCATACGACGAAGGCTCGATGAAACACACAGATGGACGGAAGATGTTCTCTATAG	1218
Db	978	CTTTTTCATAAGACGGAGGCTTGACGAACCTCCAGATGGACGCAAGACATTCGTGACAG	1037
Qy	1219	AGCTTTGCTCCATGGGCATATAGTTGAATTACTTTCGACAGCAGCAATTCCTGGAGATCTT	1278
Db	1038	AGCGTGTCTCATGGGCATATAGTTGAACTCTCCGACAGCAGCATCTCTGGAGATCTT	1097
Qy	1279	CCTGGAAGGCACAGTTTCTAGGAGTGGAAAACCTCTTGCTCGGCGAGGACTTTTGTCT	1338
Db	1098	CCTGGAAGGCACCGCTCCCGCAGTGGCAGACCTCTGTGTGCGCGGCGGGCTCCTGTC	1157
Qy	1339	AGTTTGGTATAGTACTCTGTCTACCAATGTCTATCCAGACATCTTGATATAATCTGTTTG	1398
Db	1158	AGTGGTAGTGGATACTCTGTCTATCCAAACCATCCCTGACATCTGCTGTCATCCCTGTGG	1217
Qy	1399	AATCTCTATGATCGCATTATCGAAGGTCACTACAATGGTGAACAACTGGGCAACCTAA	1458
Db	1218	CATCTGTATGATCGGATAATCGAAGGTCACTACAATGGTGAACAGTTGGGCAAGCCCA	1277
Qy	1459	GAAGAAAGAGACCTGTGGAGTGTAGCAAGAGGTGTTATTAGAAATGTTCGAAAAAACTA	1518
Db	1278	GAAGAAAGAAAGTCTCTGGAGTGTGGCAAGGCGTTATCNGAATGCTGCGGAANAATA	1337
Qy	1519	TGGTTGTGTCGAGTGGATTTTGACAGCCATTTTCTTAAGGAATATTTAGAAAGCCA	1578
Db	1338	CGGCTATGTCCGAGTGGACTTGCACAGCCATTTTCTTTGAAGGAATATTTAGAAAGCCA	1397
Qy	1579	AAGTCAGAAAACCGGTGTCTCTACTTCTTCCCTGGAGCAAGCGTGTATACAGCTATACT	1638
Db	1398	AAGTCAGAAAACCTGTATCTGCTCCCTCTCTTTTGGAGCAAGCACTGTATACAGCAATCCT	1457
Qy	1639	TCCTTCAAGACCCAGTGATGTGCTGATGAAGGTAGAGACACGCTCCATTAATGATGCAG	1698

Db 2538 AGAGACCAAGCAGAGCGGCTGTCTTAGAAGCTGAGCACCACCTTCCTACCTCAGGG 2597  
Qy 2779 CAACCGCAAAACTTCTAGATATATCTGAGTTTGTGGTGTCTGTAG 2827  
Db 2598 CAGCGGCGAGAGCTCCTGGAATACATCTGAGCTTCGTGTCTGTAG 2646

## RESULT 13

AF021348 2646 bp mRNA linear ROD 24-NOV-1999  
LOCUS Rattus norvegicus sn-glycerol-3-phosphate acyltransferase (GPAT)  
DEFINITION mRNA, nuclear gene encoding mitochondrial protein, complete cds.

ACCESSION AF021348  
VERSION AF021348.1 GI:2444458

KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus

1 (bases 1 to 2646)  
Ganesh Bhat, B., Wang, P., Kim, J. H., Black, T. M., Lewin, T. M.,  
Fiedorek, F. T. Jr. and Coleman, R. A.

TITLE Rat sn-glycerol-3-phosphate acyltransferase: molecular cloning and  
characterization of the cDNA and expressed protein  
Biochim. Biophys. Acta 1439 (3), 415-423 (1999)

REFERENCE 1 (bases 1 to 2646)  
Bhat, B. G., Wang, P., Kim, J. -H., Black, T. M., Lewin, T. M.,  
Fiedorek, F. T. Jr. and Coleman, R. A.

AUTHORS Direct Submission  
TITLE Submitted (26-AUG-1997) Nutrition, Univ. of North Carolina, CB7400  
JOURNAL McGavaran-Greenberg Bldg., Chapel Hill, NC 27599, USA

FEATURES Location/Qualifiers  
1..2646

/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/tissue\_type="liver"  
1..2646  
/gene="GPAT"  
/gene="GPAT"  
160..2646  
/EC\_number="2.3.1.15"  
/note="similar to protein encoded by Mus musculus glycerol  
3-phosphate acyltransferase gene, GenBank Accession Number  
W77003, and Accession Numbers U36771, U36772, U36773"

/codon\_start=1  
/product="sn-glycerol-3-phosphate acyltransferase"  
/protein\_id="AA071605.1"  
/db\_xref="GI:2444459"  
/translation="MEESVITGIDVSLNSSEYSLGRCKHTNEDWDVCGKPTFP  
RSATLKWESLMSKRPVGRCCCTPQSWERFNPSPISGLRNVIYINETHRR  
GWLARSLYILFVQERDVHKGFMATSIDNLSRSRVOEAIAEVAALPDGSAQOS  
KAIQKVRKAIQEMVATSPGMIRLTGWLLKLFNSFFWNIQIHKQLEMYKAAT  
ETNLPPLPVHRSHIDYLLTFILFCHNIKAPVAGNNLNIPIFSTLHLKLGFFI  
RRLEDTDPGRKIDILYRALLHGHIVELLRQQFLIELEGTSSRKTSCARGLLSV  
VDTLSSNTIPDILVPIVGSYDRIIEGHNGEQLGPKPKNESLWSVARGVIRLRKN  
YGVYVDFAPQPSLKEYLEGOSQKPSAPLSLEQALLPAILPSRPAADAAHEPDMNS  
ESRNADEAFERRILANLAELHIFTASKSCAIMSTHIVACILLYRHRQGHLSLVED  
FVNMKEVLAARDFDLFGSNGEDVVMHAIQLLNCVTITHTSRKDEFFITPSTVPSV  
FELNFSYGLVLRVIMEAIACISIVAVONKRGSGSAGGLNLIQSOLVRKAASLCY  
LLSNEGRTISLPQFYQVOCQVCKFYQGLITVAEQDDQEDVSPGLAEQWKKLPE  
TLNRSDEEDSDFEGRQDCYLKVSQAKHQFFITFLQRLGLLEAYSSAAIFVH  
TFRGPVSESYLKLHLYLLRTRNVAVYAESATYCLVKNVRFKIDIGVFKETKQ  
RASVLELSLTFPQGRQKLELYLSFVVL"  
BASE COUNT 670 a 689 c 670 g 617 t  
ORIGIN

Query Match 63.1%; Score 1896.2; DB 10; Length 2646;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2126; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 319 TTTGGAAATTACACTTTTGTGACATGGATGAATCTGCACTGACCCCTTGGTGAACAATAGATGT 378  
Db 138 TTTGGGACTTCCACGTTGAGTCATGGAGGAGTCTTCTAGTGACAAATGGCAACAATAGAGCT 197  
Qy 379 TTTCTATCTGCCACATTCATCAGAAATACAGTGTGTGTCGATGTAGAGCAACAAGTGAGGA 438  
Db 198 TTTCTATCTGCCCAATTCATCAGAAATACAGCTTTGGCCGATGTAAACACACGAATGAGGA 257  
Qy 439 ATGGGGTGAAGTGTGGCTTTAGACCCACCGTCTTCAGATCTCGAACTTAAATGGAAGA 498  
Db 258 CTGGTTGACTGTGGCTTCAAACTTCTTTCAGATCCCGAAGCTGAAATGGAAGA 317  
Qy 499 AAGCCTTAATGAGTCGGAAAGCCATTTGTTGGAGAGATGTTTACTCTCTGCACCTCCCCA 558  
Db 318 GAGCCTCATGAGCGGAAGAGGCCCTTCGTGGGAAGGTGTTGCTATTTCATGCACGCTCA 377  
Qy 559 GAGCTGGGACAAATTTTCAACCCAGTATCCCGTCTTTGGGGTTTGGCGAATGTTATTATA 618  
Db 378 GAGCTGGGAAAGGTTTTTCAACCCAGTATCCCATCTCTGGGTTTGGCGAATGTTATTATA 437  
Qy 619 TATCAATGAAACTCACACAGACACCGCGGATGCTTGCAGAGCGCTTTCTTACGTCT 678  
Db 438 TATCAATGAAACTCACACAGGACCGGAGATGCTGCGAAGCGCTTTCTTACATCT 497  
Qy 679 TTTTATTCAAGAGCGAGATGTGCAATAAGGGCATGTTTGGCCACCAATGTGACTGAAATGT 738  
Db 498 TTTTGTTCAGAGCGCGATGTCCACAGGGCATGTTTGGCCACCATGATCATCTGACAATGT 557  
Qy 739 GCTGAACAGCAGTAGATGACAGAGGCAATTCAGAGAGTGGCTGTGAATTAACCTTGA 798  
Db 558 ACTGAATAGCAGCAGAGTCCAAAGAGGCAATTTGCTGAAGTGGCTGCAGAAATTTGAACCCGGA 617  
Qy 799 TGGTCTCTCCCGCAGCAGCAATCAAAAGCCGTAAACAAGTGAAGAAAGAAAGCTAAAGAT 858  
Db 618 TGGATCTGCCCAGCAGCAGTCCAAAGCCATCCAGAAAGTGAAGAAAGAAAGCCAGGAAGAT 677  
Qy 859 TCTTCAAGAAATGTTGCCACTGTCTCACCAGCAATGATCAGACTGACTGGTGGTGGTCT 918  
Db 678 CTTCCAGGAATGTTGCTACAGTCTCCCCGGGATGATCAGGCTGACTGGCTGGGTGTT 737  
Qy 919 GCTAAACTGTTTCAACAGCTTTCTTTTGGAAATTTCAAAATTCACAAAGCTCAACTTGAGAT 978  
Db 738 ACTAAAGCTTCTCAACAGCTTCTTCTGGAACATTCAGATTCAACAGGGTCAACTTGAGAT 797  
Qy 979 GGTAAAGCTGCAACTGAGAGCAATTTGCCGCTTCTGTTTCTACAGTTCATAGATCCCA 1038  
Db 798 GGTAAAGCTGCAACTGAGAGCAATCTGCCGCTCTTGTCTGCGGCTGCAAGATCCCA 857  
Qy 1039 TATTGACTATCTGCTGCTCACTTTTCTTCTGCAATAAATCAAGCACCACATACAT 1098  
Db 858 CATGACTACTCTGCTGCTCACTTCTTCTGCAATAAATCAAGCTCAACTACAT 917  
Qy 1099 TGCTTCAGGCAATTAATCTCAACATCCCAATCTTCAGTACCTTTGATCCATAGTTGGGG 1158  
Db 918 CGCTTCGGGCAACCACTCAACATCCCATCTTTCAGTACCTTTGATTCACAAGCTTGGGG 977  
Qy 1159 CTTCTTTCATCAGCAAGGCTCGATGAAACACAGATGGAGCAAGAAAGATTTCTCTATAG 1218  
Db 978 CTTTTCATAGACGAGGCTTGACGAATCTCCAGATGGAGCAAGACATTTCTGTACAG 1037  
Qy 1219 AGCTTTCCTCCTCAGGCAATAGTTGAATTAATTCGACAGCAGCAATTTCTTGGAGATCTT 1278  
Db 1038 AGCTTTCCTCCTCAGGCAATAGTTGAATTAATTCGACAGCAGCAATTTCTTGGAGATCTT 1097  
Qy 1279 CTTGGAAGGCAAGCTTCTAGGAGTGGAAACCTCTTGTCTCGGGCAGGACATTTTGTCTC 1338  
Db 1098 CTTGGAAGGCAAGCTTCTCGGAGTGGCAAGACCTCTGTCGCCCGCGGGCTCTCTGTC 1157  
Qy 1339 AGTTTGGTGAATGACTCTGTCTTCAATGTCATCCAGACATCTTGTGATAATACCTGTGG 1398  
Db 1158 AGTGGTAGTGGACTCTGTCTCATCAACACCATCCCTGACATCTCTGTGTCATCTCTGTGGG 1217

QY 1399 AATCTCTTATGATCGATTATCGAAGGTCACTAAATGGTGAAACAATGGGCAAACTAA 1458  
Db 1218 CATCTCGTATGATCGGATTAATCGAAGTCACTAAATGGTGAAACAGTTGGGCAAGCCAA 1277  
QY 1459 GAAGAATGAGAGCCTGTGGAGTGTAGCAAGAGTGTATTTAGAAATGTTAGCAAAAACATA 1518  
Db 1278 GAAGAATGAAGTCTCTGGAGTGTGCAAGAGGCGTTATCAGAATGCTGCGGAAAACATA 1337  
QY 1519 TGGTGTGTCGAGTGGATTTTCACAGCCATTTTCTTAAAGGAATATTTAGAAAGCCA 1578  
Db 1338 CGGCTATGTCGAGTGGATTTTCACAGCCATTTTCTTAAAGGAATATTTAGAAAGCCA 1397  
QY 1579 AAGTCAGAAAACCGGTGTCTGCTCTACTTTTCCCTGGAGCAAGCGTTGTTACCACTATACT 1638  
Db 1398 AAGTCAGAAAACCGGTGTCTGCTCTACTTTTCCCTGGAGCAAGCGTTGTTACCACTATCT 1457  
QY 1639 TCCTTCAAGACCCAGTGTCTGCTGATGAGGTAGAGACAGTCATTAATGAGTCCAG 1698  
Db 1458 TCCTTCAAGACCCAGTGTCTGCTGATGAGGTAGAGACAGTCATTAATGAGTCCAG 1517  
QY 1699 AATGCAACAGATGAATCCCTAGAGAGGTTGATTGCAAAATCTGGCTGAGCATATCT 1758  
Db 1518 AAACGGCGCAGACGAAGCCCTTCGAAGAGGCTGATCGCAAACTGGCGAGCACATCT 1577  
QY 1759 ATTCACTGTCTAGCAAGTCTGTGCTCCATTATGTCACACACATTTGTGGCTTGCCTCTCT 1818  
Db 1578 CTTTACCGCCAGCAAGTCTGTGCTCCATTATGTCACACACATTTGTGGCTTGCCTCTCT 1637  
QY 1819 CTAAGACACAGCAGGGAATGATCTCTCCATGTCGAGAGACTTCTTGTGATGAA 1878  
Db 1638 CTAAGACACAGCAGGGAATGATCTCTCCATGTCGAGAGACTTCTTGTGATGAA 1697  
QY 1879 AGAGGAAGTCTTGGCTGATTTTACCTGGGCTTCTCAGGAATTTCAAGATCTAGT 1938  
Db 1698 GAGGAAGTCTTGGCTGATTTTACCTGGGCTTCTCAGGAATTTCAAGATCTAGT 1757  
QY 1939 AATGCAATGCCATACAGTGTGGGAAATTTGTGTCACAATCACCCACATAGCAGGAACGA 1998  
Db 1758 CATGCAATGCTATTTACGCTTCTGGGAACTGTGTCACAATCACCCACATAGCAGGAACGA 1817  
QY 1999 TGAGTTTTTTATACCCCGACAGCAACTGTGCCATCAGTCTTCGAACTCAGACTTCTACAG 2058  
Db 1818 TGAATTTCTTTATTAATCCAGCAACTGTGCCGCTCTTGTGAATCTCAACTTCTACAG 1877  
QY 2059 CAATGGGTACTTCATGCTTTATCATGAGGCGCATCATAGTTCGAGCCTTTATGCACT 2118  
Db 1878 CAATGGGTACTTCATGCTTTATCATGAGGCGCATCATAGTTCGAGCCTTTATGCACT 1937  
QY 2119 TCTGAACAAGAGGGGACTGGGGGTCCCACTAGCACCCCACTTAACCTGATCAGCCAGGA 2178  
Db 1938 CCAGAATAGAGGGGTTCCGAGGGTCTGCCGAGGCTTGGCAACTGATCAGCCAGGA 1997  
QY 2179 GCAGTGTGTCGGAAGCGGCCAGCTGTGCTACTTCTTCTCAATGAAGGCAACATCTC 2238  
Db 1998 GCAGTGTGTCGGAAGCGGCCAGCTGTGCTACTTCTTCTCAATGAAGGCAACATCTC 2057  
QY 2239 ACTGCTTGCACACATTTTACCAAGTCTGCCATGAACAGTAGGAAAGTTTATCCAGTA 2298  
Db 2058 TCTGCCCTGCCAGACATTTTACCAGGTTGTCAAGAGACAGTAGGAAAGTTTATCCAGTA 2117  
QY 2299 TGGCATTTCTTACAGTGGCAGACGATGACACGAGAGATATCAGTCTTGTCTGA 2358  
Db 2118 CGGAATTTCTACAGTGGCAGACGATGACACGAGAGATGTCAGTCTTGTCTGAGTA 2177  
QY 2359 GCAGAGTGGGACAGAGCTTCCAGAACCTTTGTCTTGAGAGAGTGTAGGAAGATGA 2418  
Db 2178 GCAGAGTGGGACAGAGCTTCCAGAACCTTTGTCTTGAGAGAGTGTAGGAAGATGA 2237  
QY 2419 AGACAGTGTCTTGGGAGGAAACAGCGAGATTTCTTACCTGAAGTGTAGGCAATCCAAGGA 2478  
Db 2238 GGCAGTGTCTTGGGAGGAGGAGGCTGATTTCTTACCTGAAGTGTAGGCAATCCAAGGA 2297  
QY 2479 GCACAGCAGTGTATCACCTTCTTACAGAGACTCTTGGGCTTTTGTGGAGGCTTACAG 2538

Db 2298 GCACAGCAATTCATCACCTTTCTGCAGAGGCTTCTGGGCGCCCTGCTAGAAGCCTACAG 2357  
QY 2539 CTCTGCTGCACCTTTGTTTCAAACTTCTAGTCTCTGTTTCCAGAACCTTCTGAGTATCTGCA 2598  
Db 2358 CTCTGCTGCACCTTTGTTTCAAACTTCTAGTCTCTGTTTCCAGAACCTTCTGAGTATCTGCA 2417  
QY 2599 AAAGTTGGCAAAATACCTAATAACCAAGAACAGAAATGTTTGCAGTATATCTCTGAGAG 2658  
Db 2418 GAAGCTGCACAGGTACTTCTCACAGGACGAGAGAACGTCGCGGTGTACGCTGAGAG 2477  
QY 2659 TGCACATATGCTTGTGCAAGATCTGTGTTTAAAGATGTTTAAAGATGTTTAAAGATGTTTAA 2718  
Db 2478 TGCACATATGCTTGTGCAAGATCTGTGTTTAAAGATGTTTAAAGATGTTTAAAGATGTTTAA 2537  
QY 2719 GGAGACCAACAAAGAGAGTGTCTGTTTAAAGATGTTTAAAGATGTTTAAAGATGTTTAA 2778  
Db 2538 AGAGACCAACAAAGAGAGTGTCTGTTTAAAGATGTTTAAAGATGTTTAAAGATGTTTAA 2597  
QY 2779 CAACCGACAAAATCTTCTAGATATATCTGAGTCTTGTGTTTGTGTTGTTGTTGTTGTTG 2827  
Db 2598 CAGCCGGCAGAGCTCTCTGGAATACATCTGAGCTTCTGTTGTTGTTGTTGTTGTTGTTG 2646

RESULT 14  
RNU36771  
LOCUS  
DEFINITION  
Rattus norvegicus sn-glycerol 3-phosphate acyltransferase mRNA, linear ROD 26-MAR-1999  
ACCESSION  
U36771 U36772 U36773  
VERSION  
U36771.2 GI:4481754  
KEYWORDS  
Rattus norvegicus (Norway rat)  
SOURCE  
Rattus norvegicus  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
1 (bases 1 to 2519)  
AUTHORS  
Nikonov, A.V., Morimoto, T. and Haldar, D.  
TITLE  
Properties, purification and cloning of mitochondrial sn-glycerol  
3-phosphate acyltransferase  
(in) Pandalai, S.G. (Ed.);  
JOURNAL  
RECENT RESEARCH DEVELOPMENTS IN LIPIDS RESEARCH, VOLUME 2: 207-222;  
Transworld Research Network, India (1998)  
REFERENCE  
2 (bases 1 to 2519)  
AUTHORS  
Nikonov, A.V., Morimoto, T. and Haldar, D.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (20-SEP-1995) Andrei V. Nikonov, Department of Biological  
Science, St. John's University, Jamaica, NY 11439, USA  
REFERENCE  
3 (bases 1 to 2519)  
AUTHORS  
Nikonov, A.V., Morimoto, T. and Haldar, D.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (23-MAR-1999) Cell Biology, New York University School of  
Medicine, 550 1st Ave. MSB-697, New York, NY 10016, USA  
REMARK  
Sequence update by submitter  
COMMENT  
On or before Mar 23, 1999 this sequence version replaced  
gi:1754786, gi:1754788, gi:1754784.  
FEATURES  
Location/Qualifiers  
1..2519  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague Dawley"  
/db\_xref="taxon:10116"  
/sex="male"  
/tissue\_type="liver"  
/clone\_lib="Stratage cat. number 936513"  
101..2419  
/EC\_number="2.3.1.15"  
/note="similar to product encoded by Mus musculus glycerol  
3-phosphate acyltransferase gene, GenBank Accession Number  
M77003"  
/codon\_start=1  
/product="sn-glycerol 3-phosphate acyltransferase"

```

/protein_id="AAB39470.2"
/db_xref="GI:4481755"
/translation="MSRRKPFVGRCCYSCSTPQSWERFPNPSIPPLGLRNVIYINETH
RQGWLARLSYILFVDRDVHKGWFATSIIDNVLSRVQEAIAEVAALVDPGSAQ
QSKAIQKVKARKILQEMVATVSPGMIRLTQVLLKLFNSFFNIIQIHKGOLEMVK
AATNPLPLPVHRSHIDYLLATFLLFCHNIKAPYIASNNLNLIPIFSLIHLGG
FFIRRLDTPDGRKIDILYRALLGHVVVELLQQOFLIFLEGFRSRGKTSRAGV
LSVVDTLSNTIIPDILIPVIGISYDRILIEGHYNGEOLGPKKNESLWSVARGVIRL
KNGYVVRVDFAPRSLKEYLEQGSQKPSVAPLSLEQALLPAILPSRPAADAAEHDM
SSNRSNAADFARRKLIJANLAENHLLTASCSAIMTHIVACILLYRHWOIHLSTL
VEDFVKEEVLNDFDLGRFSGNSEDVVMHAIQLLGNCTVITHTSRKDERPITPSTTV
PSVPELNFYSGLVDFVMEALJACSIYAVONKRGSGGLPEPGLNISQEQVLVRAASL
CYLLSNEGTISLPQCTFQVQCHETVGFQIGILTVAEQDDQEDVDSGLAEQQMNKL
PEPLNWRSDDEDSDFGEQRDCYLKVSQAKHEQOFTFLORLLGLLEAYSAATP
HTFRAVPVPESEYIQLKHRHYLLTRTERNAVAYAESATVCLVKNVAKMFKDITGVFKETK
QKRASVLELS"TFLPQSRQKXLEVLISFVL"
transit_peptide 101..286
misc_feature 1367..1405
/evidence=not_experimental
misc_feature 1658..1705
/evidence=not_experimental
misc_feature 254..286
/note="encodes R(-10) motif recognizable by matrix
proteases"
BASE COUNT 627 a 670 c 644 g 578 t
ORIGIN
Query Match 60.2%; Score 1807.2; DB 10; Length 2519;
Best Local Similarity 84.2%; Pred. NO. 0;
Matches 2050; Conservative 0; Mismatches 383; Indels 3; Gaps 1;
QY 406 CAGTGTTCGATGATGACACACAAAGTGGGATGGGTGAGTGTGCTTAGACCCAC 465
DB 1 CAGCGTTCGCGATGATAACACACAAATGAGGACTGGGTGACTTTGGCTTCAAACTTAC 60
QY 466 CGTCTTCAGATCTGCAACTTTAAATGGAAGAAGCCCTAATGATCGGAAGAGCCATT 525
DB 61 CTTCTTCAGATCCGCAAGCTGAAATGGAAGAGAGGCTCATGAGCCGGAAGAGCCCTT 120
QY 526 TGTGGAAAGATGTTGTTACTCTGTCACATCCCGAGAGTGGGACAAATTTTCAACCCGAG 585
DB 121 CGTGGGAAGGTGTTGCTATTATGACGACGCTCAGAGCTGGGAAAGGTTCTTCAACCCGAG 180
QY 586 TATCCCGTCTTGGGTTTGGGATGTTTATTTATATCAATGAACTCACAGACACCG 645
DB 181 TATCCCACTCTGGGTTTGGGAAATGTTTATTTATATCAATGAACTCACAGAGCACCG 240
QY 646 CGGATGGCTTCAAGACGCTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTGCATAA 705
DB 241 AGGATGGCTGCAAGACGCTTTCTTACATCTTTTGTTCAGAGCCGATGTCCACAA 300
QY 706 GGGCATGTTTCCCAACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAGAGGC 765
DB 301 GGGCATGTTTCCCAACCAATGTGACTGAAATGTGCTGAATAGCAGCAGTCCAGAGGC 360
QY 766 AATTGCAGAGTGGCTGCTGTAATTAACCTGTAGTGTTCGCCAGCAGCAATCAAAAGC 825
DB 361 AATTGCTGAAGTGGCTGAGAAATTAACCCGGATGGATCTGCCAGCAGCAGTCCAAAGC 420
QY 826 CGTTAAACAAAGTGAAGAAGCTTAAAGATTTCTTCAAGAAATGGTTGCCACTGTCTC 885
DB 421 CATCAGAAATGAAGAAGCCAGGAAGATCTCCAGGAATGGTTGTGTCAGTCTC 480
QY 886 ACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTGCTAAACCTGTTTCAACAGCTTCTTTTG 945
DB 481 CCCCAGGATGATCAGGCTGACTGGCTGGGTGTACTAAAGCTCTTCAACAGCTTCTTCTG 540
QY 946 GAACATTCAAATTCACAAAGTCAACTTGGAGATGGTTAAAGCTCAACTGAGACGAATTT 1005
DB 541 GAACATTCAGATTCACAGGGTCAACTTGGAGATGGTGAAGCTGCAACTGAGACGAATCT 600

```

```

QY 1006 GCCGCTTCTGTTTCTACCAAGTTCTATAGATCCCATATTGACTATCTGCTGCTCACTTTCAT 1065
DB 601 GCCGCTCTGTTTCTGCGGGTGACAGATCCCATCGACTACCTGTTGCTCACTTTCAT 660
QY 1066 TCTCTTCTGCCATTAACATCAAGCACCATACATTCGTTTCAGGCAATATCTCAACATCCC 1125
DB 661 CCTCTTCTGCCAACAACTCAAGCTCCATATACGCTCGGCGCAACAACTTCAACATCCC 720
QY 1126 AATCTTCAGTACCTTGTATCCATAAGCTTGGGGGCTTCTTATACAGCAGAGGCTCGATGA 1185
DB 721 CATCTTCAGTACCTTGTATTCACAAGCTTGGGGGCTTTTTTATAAGACGGAGGCTTGACGA 780
QY 1186 AACACCAAGATGGACGGAAAGATGTTCTCTATAGAGCTTTTGTCTCATGGGCAATATAGTTGA 1245
DB 781 AACTCCAGATGGACGCAAGACATTTCTGTACAGAGCGTTGCTCCATGGGCATGTAGTTGA 840
QY 1246 ATTACTTCGACAGCAGCAATTTCTTGGAGATCTTCTGGAAGGCACAGCTTCTAGAGTGG 1305
DB 841 ACTCTCCGACAGCAGCAGTTCCTTGGAGATCTTCTGGAAGGCACCGCTCCCGCAGTGG 900
QY 1306 AAAAACCTCTTGTCTCGGCGCAGCACTTTTGTCAAGTTGTGTAGATATCTCTGTCTACCAA 1365
DB 901 CAGACCTCTCTGTGCCCGCGCGCTCCTGTCAAGTGTGTAGTGTGATATCTGTCTATCCAA 960
QY 1366 TGTCAATCCCAAGACATTTGATATACCTGTGTGAATCTCTTATGATCGCATATTCGAAGG 1425
DB 961 CACCATCCCTGACATCTGCTGTCATCTCTGTGGGATCTCTGTATGATCGGATATTCGAAG 1020
QY 1426 TCATCAATAGTGTGAACAACTGGGCAAACTTAAGAAGATGAGAGCCTGTGGAGTGTAGC 1485
DB 1021 TCATCAATAGTGTGAACAACTGGGCAAACTTAAGAAGATGAGAGTGTGTGGAGTGTGGC 1080
QY 1486 AAGAGGTGTATTAAGATTTACGAAATACTATGTTGTGTGTCGAGTGGATTTTGACA 1545
DB 1081 AAGAGGCTTATCAGAATGTGCGGAAATACTACGGCTGTGTGTCGAGTGGACTTTTGACA 1140
QY 1546 GCCATTTTCTTAAAGAAATATTTAGAAAGCCAAAGTCAGAAACGGTGTCTGCTCTACT 1605
DB 1141 GCCATTTTCTTGAAGAAATATTTAGAAAGCCAAAGTCAGAAACCTGTATCTGCTCCCT 1200
QY 1606 TTCCCTCGAGCAAGCTGTTTATCCAGCTATATCTTCTTCAAGACCCAGTGTGCTGTGA 1665
DB 1201 CTCTTTGGAGCAAGCACTGTTACCGCAATCTTCTTCAAGACCTGATGCTGTGCTGC 1260
QY 1666 TGAAGTGAAGACACATGTCCTAATGATCCAGAAATGCAACAGATGAATCCCTACGAG 1725
DB 1261 CGAACATGAAGACATGTCCTAATGATCCAGAAACCGCGCAGACGAAAGCTTTCCGAAG 1320
QY 1726 GAGGTTGATTTGCAATCTGCTGAGCATATTTATTTCACTGTCTAGCAAGTCTCTGCGCAT 1785
DB 1321 GAGGCTGATGCAAACTGCGGAGCACAATCTCTCACCGCCAGCAAGTCTCTGCGCTAT 1380
QY 1786 TATGTCACACACATTTGCTGCTGCTCTCTACAGACACAGCAGGAGGAAATGATCT 1845
DB 1381 CATGTCCACCACTTGTGCGCTGCTCTCTACAGACACCTGCGCAGGAGAAATCCACCT 1440
QY 1846 CTCCACATTTGGTTCGAAAGCTTTTGTGATGAAGAGAGAGTCTCTGGCTCGTGAATTTGA 1905
DB 1441 CTCCACGCTGGTGGAGACTTTTGTGATGAAGAGAGAGTCTCTAGCTCGGAGATTTGA 1500
QY 1906 CTTGGGCTTCTCAGGAAATTCAGAAAGATGTAGTAATCATGCCCATACAGCTGCTGGGAAA 1965
DB 1501 CTTGGGCTTCTCAGGAAATTCAGAAAGATGTAGTAATCATGCCCATACAGCTTCTGGGCAA 1560
QY 1966 TTGTGTCACAACTACCCACACTAGCAGGAAAGTGTATTTTATACCCCCCAGCAAC 2025
DB 1561 CTGTGTCACAACTACCCACACTAGCAGGAAAGTGTATTTTATTTACTCCACACACAC 1620
QY 2026 TGTCCCATCAGTCTTGGAACTCAACTTCTACAGCAATGGGGTACTTCTCATGTCTTTATCAT 2085
DB 1621 TGTCCCGCTCGTCTTGTGAATCTCAACTTCTACAGCAATGGGGTACTTCTCATGTCTTTATCAT 1680

```



Result No.	Score	Query Match	Length	DB	ID	Description
1	3001	99.9	3003	24	ABK11094	cDNA encoding huma
2	2970.2	98.9	4031	23	ABV25313	Human prostate exp
3	2958	98.5	3210	24	AAD35221	Human TRNFR-6 cDNA
4	2591.4	86.3	2684	25	ABX72233	Human NOVX polymuc
5	2485.4	82.8	2487	24	ABK94821	DNA encoding human
6	2483.8	82.7	2487	25	AAL55475	GPAM related DNA s
7	2334	77.7	3273	24	ABZ12025	Human polynucleoti
8	1896.2	63.1	2646	24	ABK63741	Rat sequence diffe







[illegible]



CC for treating disorders associated with a decreased expression of  
CC functional TRNFR, e.g., cell proliferative, developmental, neurological,  
CC autoimmune/inflammatory disorders and parasitic infections. Antagonists  
CC of TRNFR proteins are useful in treating disorders associated with  
XX increased activity of TRNFR. The present sequence is human TRNFR-6 cDNA.

SQ Sequence 3210 BP; 896 A; 721 C; 735 G; 858 T; 0 other;  
Query Match 98.5%; Score 2958; DB 24; Length 3210;  
Best Local Similarity 99.74; Pred. No. 0;  
Matches 2983; Conservative 0; Mismatches 7; Indels 2; Gaps 2;  
QY 12 GCTGCTCGCGGGGAGCTCTTCTGAGGTACTGTGGAGCACCCAAAGTCTGTGAGCTCT 71  
DB 9 GCTGCTCGCGGGGAGCTCTTCTGAGGTACTGTGGAGCACCCAAAGTCTGTGAGCTCT 68  
QY 72 GCGCGTCAACAGGCAACCCAGAGGAACAGAGCTTGTCTTATTCACCCACAGCTGGGAC 131  
DB 69 GCGCGTCAACAGGCAACCCAGAGGAACAGAGCTTGTCTTATTCACCCACAGCTGGGAC 128  
QY 132 TGTCTTCTCCAGAGTCTCCATCAGCTTTGCTTAATCGACTGATTTGGAATAATTTCTCMAA 191  
DB 129 TGTCTTCTCCAGAGTCTCCATCAGCTTTGCTTAATCGACTGATTTGGAATAATTTCTCMAA 188  
QY 192 CACCACCAAGTCAAGGATACAGGAGGAGGAGGCTCCCTCTGTGTATGAGCAATTCGCAAC 251  
DB 189 CACCACCAAGTCAAGGATACAGGAGGAGGAGGCTCCCTCTGTGTATGAGCAATTCGCAAC 248  
QY 252 CGAAACTGATAGTGTAGTCTGAAAGTTTATGTTATGAACAGAGAACTTTTATCCAG 311  
DB 249 CGAAACTGATAGTGTAGTCTGAAAGTTTATGTTATGAACAGAGAACTTTTATCCAG 308  
QY 312 CACATGATTTGGAAATTTACATTTGTGACATGATGATGATGATGATGATGATGATGATGAT 371  
DB 309 CACATGATTTGGAAATTTACATTTGTGACATGATGATGATGATGATGATGATGATGATGAT 368  
QY 372 TAGATGTTTCTTATCTGCCACATTCATCAGAAATACAGTGTGTGATGATGATGATGATGATGAT 431  
DB 369 TAGATGTTTCTTATCTGCCACATTCATCAGAAATACAGTGTGTGATGATGATGATGATGATGAT 428  
QY 432 GTGAGGAATGGGTGATGTGGCTTTAGACCCACCGTCTTCAGATCTGCAACTTTAAAT 491  
DB 429 GTGAGGAATGGGTGATGTGGCTTTAGACCCACCGTCTTCAGATCTGCAACTTTAAAT 488  
QY 492 GGAAGAAAGCTTAATGAGTCGGAAGAGGCAATTTGTGGAAGATGTTGTACTCTGCA 551  
DB 489 GGAAGAAAGCTTAATGAGTCGGAAGAGGCAATTTGTGGAAGATGTTGTACTCTGCA 548  
QY 552 CTCCCAGAGCTGGGACAAATTTTCAACCCAGTATCCCGTCTTTGGGTTTGGGAATG 611  
DB 549 CTCCCAGAGCTGGGACAAATTTTCAACCCAGTATCCCGTCTTTGGGTTTGGGAATG 608  
QY 612 TTAATTTATCAATGAATCTCACAAAGACCCGCGATGGCTTGTGAAGACGCTTTCTT 671  
DB 609 TTAATTTATCAATGAATCTCACAAAGACCCGCGATGGCTTGTGAAGACGCTTTCTT 668  
QY 672 ACCTTCTTTTATTCAGAGGAGATGTGATAGGCAATTTTGCACCAATGTGACTG 731  
DB 669 ACCTTCTTTTATTCAGAGGAGATGTGATAGGCAATTTTGCACCAATGTGACTG 728  
QY 732 AAAATGTCTGAACAGAGTAGAGTACAAGAGGCAATTTGAGAGTGGCTGTGCAATTA 791  
DB 729 AAAATGTCTGAACAGAGTAGAGTACAAGAGGCAATTTGAGAGTGGCTGTGCAATTA 788  
QY 792 ACCCTGATGTTCTGCCAGCAGCAATCAAAAGCCGTTAAACAAAGTGAAGAAAGAGCTA 851  
DB 789 ACCCTGATGTTCTGCCAGCAGCAATCAAAAGCCGTTAAACAAAGTGAAGAAAGAGCTA 848  
QY 852 AAAGGATCTTCAAGAAATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGAGGT 911  
DB 849 AAAGGATCTTCAAGAAATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGAGGT 908  
QY 912 GGGTGTCTTAAACTGTTCAACAGCTTCTTTTGGAAATTTCAAAATTCACAAAGGTCAAC 971

DB 909 CGGTGCTGCTAAAACTGTTTCAACAGCTTCTTTTGGAAATTTCAAAATTCACAAAGGTCAAC 968  
QY 972 TTGAGATGTTTAAAGCTGCAACTGAGACGAATTTGCCGCTTCTGTTTCTTACAGTTTCATA 1031  
DB 969 TTGAGATGTTTAAAGCTGCAACTGAGACGAATTTGCCGCTTCTGTTTCTTACAGTTTCATA 1028  
QY 1032 GATCCCATATTCAGTATCTGCTGCTCAGTTCATTTCTTCTTCCCATACATCAAAAGCAC 1091  
DB 1029 GATCCCATATTCAGTATCTGCTGCTCAGTTCATTTCTTCTTCCCATACATCAAAAGCAC 1088  
QY 1092 CATACATTTCTTTCAGGCAATAATCTCAACATCCCAATCTTTCAGTACCTTGTATCCATAAGC 1151  
DB 1089 CATACATTTCTTTCAGGCAATAATCTCAACATCCCAATCTTTCAGTACCTTGTATCCATAAGC 1148  
QY 1152 TTGGGGGCTTCTTCATACGACGAAGCTCGATGAACACACAGATGGACGGAAGATGTTTC 1211  
DB 1149 TTGGGGGCTTCTTCATACGACGAAGCTCGATGAACACACAGATGGACGGAAGATGTTTC 1208  
QY 1212 TCTATAGAGCTTTGCTTCCATGGGCATATAGTTGAATTTACTTCCAGACAGCAATTTCTTGG 1271  
DB 1209 TCTATAGAGCTTTGCTTCCATGGGCATATAGTTGAATTTACTTCCAGACAGCAATTTCTTGG 1268  
QY 1272 AGATCTTCTCGGAAGGCAACAGCTTCTAGAGTGGAAAAACCTTCTTGTCTCGGGCAGGAC 1331  
DB 1269 AGATCTTCTCGGAAGGCAACAGCTTCTAGAGTGGAAAAACCTTCTTGTCTCGGGCAGGAC 1328  
QY 1332 TTTTGTGAGTGTGTGTAGATCTCTGTCTACCAATGTCTCCAGACATCTTTGATATATAC 1391  
DB 1329 TTTTGTGAGTGTGTGTAGATCTCTGTCTACCAATGTCTCCAGACATCTTTGATATATAC 1388  
QY 1392 CTGTTGGAATCTCTATGATCGCATTTATCGAAGGTCTACTTACCAATGGTGAACAACTGGGCA 1451  
DB 1389 CTGTTGGAATCTCTATGATCGCATTTATCGAAGGTCTACTTACCAATGGTGAACAACTGGGCA 1448  
QY 1452 AACCTAAGAAAGTGAAGAGCTGTGAGTGTAGCAAGAGGTGTATTAGAATGTTTACGAA 1511  
DB 1449 AACCTAAGAAAGTGAAGAGCTGTGAGTGTAGCAAGAGGTGTATTAGAATGTTTACGAA 1508  
QY 1512 AAAATCTATGTTGTGTGCGAGTGGATTTTGCACAGGCAATTTTCTTAAAGGAATATTAG 1571  
DB 1509 AAAATCTATGTTGTGTGCGAGTGGATTTTGCACAGGCAATTTTCTTAAAGGAATATTAG 1568  
QY 1572 AAAGCCAAAGTCAGAAACCGGTGTCTCTACTTTTCCCTGGAGCAAGCGTGTGTACCAAG 1631  
DB 1569 AAAGCCAAAGTCAGAAACCGGTGTCTCTACTTTCTTCCCTGGAGCAAGCGTGTGTACCAAG 1628  
QY 1632 CTATATCTTCTTCAAGACCCAGTGTGCTGTGATGAAGGTAGAGACACGTCATTAATG 1691  
DB 1629 CTATATCTTCTTCAAGACCCAGTGTGCTGTGATGAAGGTAGAGACACGTCATTAATG 1688  
QY 1692 AGTCCAGAAATGCAACAGATGAATCTCTACGAGGAGGTGTGATGCAAAATCTGGCTGAGC 1751  
DB 1689 AGTCCAGAAATGCAACAGATGAATCTCTACGAGGAGGTGTGATGCAAAATCTGGCTGAGC 1748  
QY 1752 ATATTTCTATTTACTGTCTAGCAAGTCTGTCGCAATTTGTCTCCACATTTGGTCGAAGACTTCTTTG 1871  
DB 1749 ATATTTCTATTTACTGTCTAGCAAGTCTGTCGCAATTTGTCTCCACATTTGGTCGAAGACTTCTTTG 1868  
QY 1812 TGTCTCTCTACAGACACAGGCGAGGAATTTGATCTCTCCACATTTGGTCGAAGACTTCTTTG 1871  
DB 1809 TGTCTCTCTACAGACACAGGCGAGGAATTTGATCTCTCCACATTTGGTCGAAGACTTCTTTG 1868  
QY 1872 TGAAGAAAGAGGAGTCTTGGCTGTGATTTTGAACCTGGGGTCTCAGGAAATTCAGAAG 1931  
DB 1869 TGAAGAAAGAGGAGTCTTGGCTGTGATTTTGAACCTGGGGTCTCAGGAAATTCAGAAG 1928  
QY 1932 ATGTAGTAATGATGCATACAGTCTCGGAAATTTGTGTCAAAATCCACCACTAGCA 1991  
DB 1929 ATGTAGTAATGATGCATACAGTCTCGGAAATTTGTGTCAAAATCCACCACTAGCA 1988  
QY 1992 GGAACGATGAGTTTTTTTATCACCCCGACCAACGTCTCCCATCAGTCTTCGAACTCAACT 2051

Db 1989 GGAACGATGAGTTTATATACACCCCGACGACAACTGTCCATCACTCTTGGAACTCAACT 2048  
QY 2052 TCTACAGCAATGGGTACTTTCATGTCTTTATCATGGAGCCCATCATAGCTTGCAGCCTTT 2111  
Db 2049 TCTACAGCAATGGGTACTTTCATGTCTTTATCATGGAGCCCATCATAGCTTGCAGCCTTT 2108  
QY 2112 ATGCAAGTTCTGAACAAAGAGGAGCTGGGGGTCCTACCTAGACCCCACTTAACCTGATCA 2171  
Db 2109 ATGCAAGTTCTGAACAAAGAGGAGCTGGGGGTCCTACCTAGACCCCACTTAACCTGATCA 2168  
QY 2172 GCCAGGACGAGCTGTGTGGGAAGGCGGCAGCGCTGTCTTCTCTCCAAATGAAGCA 2231  
Db 2169 GCCAGGACGAGCTGTGTGGGAAGGCGGCAGCGCTGTCTTCTCTCCAAATGAAGCA 2228  
QY 2232 CCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAACACAGTAGGAAAGTTTA 2291  
Db 2229 CCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAACACAGTAGGAAAGTTTA 2288  
QY 2292 TCCAGTATGGCATTTCTTACAGTGGCAGAGCAGATGACAGGAAGATATCAGTCTTAGTC 2351  
Db 2289 TCCAGTATGGCATTTCTTACAGTGGCAGAGCAGATGACAGGAAGATATCAGTCTTAGTC 2348  
QY 2352 TTGCTGAGCAGCAGTGGACAGAAAGCTTCCAGAACCTTTGTCTTGGAGAAAGTGTATGAG 2411  
Db 2349 TTGCTGAGCAGCAGTGGACAGAAAGCTTCCCTGAACCTTTGTCTTGGAGAAAGTGTATGAG 2408  
QY 2412 AAGATGAAGACAGTGAAGTGGGAGGAGAACAGCAGAGATTGTCTGAAGGTGAGCCAAAT 2471  
Db 2409 AAGATGAAGACAGTGAAGTGGGAGGAGAACAGCAGAGATTGTCTGAAGGTGAGCCAAAT 2468  
QY 2472 CCAAGGAGCACAGCAGTTTATCACTTCTTACAGAGACTCTTGGGCTTTGTCTGGAGG 2531  
Db 2469 CCAAGGAGCACAGCAGTTTATCACTTCTTACAGAGACTCTTGGGCTTTGTCTGGAGG 2528  
QY 2532 CCTCAGCTCTGCTGCCATCTTTGTCTCAACTTCAGTGGTCTCTTCCAGAACCTTAGT 2591  
Db 2529 CCTCAGCTCTGCTGCCATCTTTGTCTCAACTTCAGTGGTCTCTTCCAGAACCTTAGT 2588  
QY 2592 ATCTCAAAAGTTGCAAAATACCTTAATAACAGAAACAGAAAGAAATGTTGAGTATATG 2651  
Db 2589 ATCTCAAAAGTTGCAAAATACCTTAATAACAGAAACAGAAAGAAATGTTGAGTATATG 2648  
QY 2652 CTGAGAGTGCACATATTTGTCTGTGAAGATGTCTGTGAAGATGTTTAAAGATATTTGGGG 2711  
Db 2649 CTGAGAGTGCACATATTTGTCTGTGAAGATGTCTGTGAAGATGTTTAAAGATATTTGGGG 2708  
QY 2712 TTTTCAAGGAGCACAACAAAGAGAGTGTCTGTTTAAAGTGTGAGTGTGAGTGTCTTCTAC 2771  
Db 2709 TTTTCAAGGAGCACAACAAAGAGAGTGTCTGTTTAAAGTGTGAGTGTGAGTGTCTTCTAC 2768  
QY 2772 CTCATGCAACCGACAAAACCTTCTAGAAATATATCTGAGTTTGTGCTGTAGGTAA 2831  
Db 2769 CTCATGCAACCGACAAAACCTTCTAGAAATATATCTGAGTTTGTGCTGTAGGTAA 2828  
QY 2832 CGTGTGCACTGCTGGCAATGAAGGTATGAGATGAGTTCCTGTAGGTACCAAGCTTCT 2891  
Db 2829 CGTGTGCACTGCTGGCAATGAAGGTATGAGATGAGTTCCTGTAGGTACCAAGCTTCT 2888  
QY 2892 GGCTCAAGGTTTGAAGTGTCTTCCAGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCT 2951  
Db 2889 GGCTCAAGG-TTGAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2946  
QY 2952 CCTGGAAGCAAGTGCCTTCTTCCCTCCATGAGTCTGAGATCTTCCAGCTT 3003  
Db 2947 CCTGGAAGCAAGTGCCTTCTTCCCTCCATGAGTCTGAGTCTTCCAGCTT 2998

## RESULT 4

ABX72233

ID ABX72233 standard; cDNA; 2684 BP.

XX AC

XX ABX72233;

XX

DT 03-JUN-2003 (first entry)  
XX Human NOVX polynucleotide #64.  
XX  
KW Human; NOVX; gene; ss; metabolic disorder; cardiomyopathy; diabetes; ASD;  
hyperextension; congenital heart defect; aortic stenosis; valve disease;  
atrial septal defect; atrioventricular canal defect; ductus arteriosus;  
pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;  
tuberosus sclerosis; neuroderma; atherosclerosis; infectious disease;  
obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;  
Parkinson's disease; immune disorder; haematopoietic disorder;  
haemophilia; hypercoagulation; Crohn's disease; cancer.  
XX  
OS Homo sapiens.  
PN WO200281498-A2.  
XX  
XX 17-OCT-2002.  
PD  
XX  
PF  
XX  
XX 03-APR-2002; 2002WO-US10780.  
XX  
XX 03-APR-2001; 2001US-281086P.  
PR 03-APR-2001; 2001US-281136P.  
PR 05-APR-2001; 2001US-281863P.  
PR 06-APR-2001; 2001US-281906P.  
PR 10-APR-2001; 2001US-282020P.  
PR 10-APR-2001; 2001US-282930P.  
PR 12-APR-2001; 2001US-282934P.  
PR 12-APR-2001; 2001US-283512P.  
PR 13-APR-2001; 2001US-283710P.  
PR 17-APR-2001; 2001US-284234P.  
PR 19-APR-2001; 2001US-285325P.  
PR 20-APR-2001; 2001US-285381P.  
PR 20-APR-2001; 2001US-285609P.  
PR 23-APR-2001; 2001US-285748P.  
PR 23-APR-2001; 2001US-285890P.  
PR 24-APR-2001; 2001US-286068P.  
PR 25-APR-2001; 2001US-286292P.  
PR 27-APR-2001; 2001US-287213P.  
PR 02-MAY-2001; 2001US-288257P.  
PR 29-MAY-2001; 2001US-294164P.  
PR 30-MAY-2001; 2001US-294484P.  
PR 18-JUN-2001; 2001US-298952P.  
PR 19-JUN-2001; 2001US-299237P.  
PR 19-JUN-2001; 2001US-299237P.  
PR 12-SEP-2001; 2001US-299276P.  
PR 12-SEP-2001; 2001US-318750P.  
PR 25-SEP-2001; 2001US-324800P.  
PR 25-SEP-2001; 2001US-324802P.  
PR 27-SEP-2001; 2001US-325684P.  
PR 17-OCT-2001; 2001US-330143P.  
PR 14-NOV-2001; 2001US-332131P.  
PR 14-NOV-2001; 2001US-332240P.  
PR 14-NOV-2001; 2001US-332779P.  
PR 21-NOV-2001; 2001US-332115P.  
PR 04-DEC-2001; 2001US-337621P.  
PR 03-JAN-2002; 2002US-345783P.  
PR 16-JAN-2002; 2002US-350251P.  
PR 02-APR-2002; 2002US-0114270.  
XX  
XX (CURA-) CURAGEN CORP.  
PA  
XX  
PI Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA, Patturajan M;  
Liu X, Gusev VV, Li L, Vernet CM, Zerhusen BD, Gorman L;  
Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;  
Padigaru M, Shinkets RA, Gangolli EA, Taupier RJ, Casman SJ, Ji W;  
Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;  
Macdougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;  
Ellerman K;  
XX  
XX WPI; 2003-046858/04.  
DR P-ESDB; ABU54605.  
XX  
XX  
XX New isolated NOVX polypeptide useful for treating atherosclerosis,

PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,  
XX neurodegenerative disorders, Alzheimer's disease and cancer  
PS Claim 17; Page 231-232; 666pp; English.

XX The invention relates to human polypeptides, termed NOVX, and the  
CC polynucleotides encoding them. The polypeptides and polynucleotides are  
CC useful for diagnosing disease, and screening for potential therapeutic  
CC agents. The sequences are useful for treating metabolic disorders, aortic  
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic  
CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,  
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular  
CC septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,  
CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative  
CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,  
CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease  
CC and cancer. Sequences ABX7170-ABX72275 represent human NOVX  
XX polynucleotides of the invention.

SQ Sequence 2684 BP; 754 A; 583 C; 615 G; 732 T; 0 other;

Query Match 86.3%; Score 2591.4; DB 25; Length 2684;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2606; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY	310	AGCATGATTTGGGAATTACACTTTCTGTGACATGGATGAATCTGCACTGACCCCTTGGTAC	369
DB	70	AGCATGATTTGGGAATTACACTTTCTGTGACATGGATGAATCTGCACTGACCCCTTGGTAC	129
QY	370	AATAGATGTTTCTTATCTGCCACATTCATCAGAAATCAGATGTTGGTGCATGTAAGCACAC	429
DB	130	AATAGATGTTTCTTATCTGCCACATTCATCAGAAATCAGATGTTGGTGCATGTAAGCACAC	189
QY	430	AAGTGAAGATGGGTGAGTGGCTTTTACACCCACCGTCTTCAGATCTGCAACTTTAAA	489
DB	190	AAGTGAAGATGGGTGAGTGGCTTTTACACCCACCGTCTTCAGATCTGCAACTTTAAA	249
QY	490	ATGGAAGAAAGCCTAATGATGTCGAAAGCCGCAATTTGTTGGAGATGTTGTTACTCCTG	549
DB	250	ATGGAAGAAAGCCTAATGATGTCGAAAGCCGCAATTTGTTGGAGATGTTGTTACTCCTG	309
QY	550	CACCTCCCGAGCTGGGACAAATTTTCAACCCCGATATCCCGTCTTTGGGTTTGGGAA	609
DB	310	CACCTCCCGAGCTGGGACAAATTTTCAACCCCGATATCCCGTCTTTGGGTTTGGGAA	369
QY	610	TGTTATTTATATCAATGAAACTCACACAGACACCGCGGATGGCTTGCAAGACGCGCTTTC	669
DB	370	TGTTATTTATATCAATGAAACTCACACAGACACCGCGGATGGCTTGCAAGACGCGCTTTC	429
QY	670	TTAGGTTCTTTTATTTCAAGAGCGAGATGTCATAGGGCATGTTTCCACCAATGTGAC	729
DB	430	TTAGGTTCTTTTATTTCAAGAGCGAGATGTCATAGGGCATGTTTCCACCAATGTGAC	489
QY	730	TGAAATGTGCTCAACAGCAGTAGATGACAGAGGCAATTCAGAGTGGCTGCTGAAT	789
DB	490	TGAAATGTGCTCAACAGCAGTAGATGACAGAGGCAATTCAGAGTGGCTGCTGAAT	549
QY	790	AAACCTGTGTTTCTCCCGACGACAAATCAAAAGCCGTTAAACAAAGTGAAGAAAGC	849
DB	550	AAACCTGTGTTTCTCCCGACGACAAATCAAAAGCCGTTAAACAAAGTGAAGAAAGC	609
QY	850	TAAAGATTTCTTCAAGAAATGGTGGCATGTCATACCGGCATGATCAGATGACTGG	909
DB	610	TAAAGATTTCTTCAAGAAATGGTGGCATGTCATACCGGCATGATCAGATGACTGG	669
QY	910	GTGGGTGCTGTAAAACCTGTTTCAACAGCTTTTGGAAATTCATAATTCACAAAGTCA	969
DB	670	GTGGGTGCTGTAAAACCTGTTTCAACAGCTTTTGGAAATTCATAATTCACAAAGTCA	729
QY	970	ACTTGAGATGGTTAAAGCTGACAGCAATTTGCCGCTTCTGTTTCTACAGTTCA	1029
DB	730	ACTTGAGATGGTTAAAGCTGACAGCAATTTGCCGCTTCTGTTTCTACAGTTCA	789

QY	1030	TAGATCCCATATTGACTATCTCTGCTCACTTTCAATCTCTTCTGCCATAACAATCAAGC	1089
DB	790	TAGATCCCATATTGACTATCTCTGCTCACTTTCAATCTCTTCTGCCATAACAATCAAGC	849
QY	1090	ACCATACATTGCTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTGATCCATAA	1149
DB	850	ACCATACATTGCTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTGATCCATAA	909
QY	1150	GCTTGGGGCTTCTTCATACGACGAAGGCTCGATGAACACACAGATGGACGGAAGATGT	1209
DB	910	GCTTGGGGCTTCTTCATACGACGAAGGCTCGATGAACACACAGATGGACGGAAGATGT	969
QY	1210	TCTCTATAGAGCTTGTCTCCATGGGCATATAGTTGAATTAATCTTCGACAGCAATTCCT	1269
DB	970	TCTCTATAGAGCTTGTCTCCATGGGCATATAGTTGAATTAATCTTCGACAGCAATTCCT	1029
QY	1270	GGAGATCTTCTCGAAGGACACAGTCTTAGAGTGGAAAAACCTCTTGTCTCGGCGAGG	1329
DB	1030	GGAGATCTTCTCGAAGGACACAGTCTTAGAGTGGAAAAACCTCTTGTCTCGGCGAGG	1089
QY	1330	ACTTTTGTGAGTGTGGTAGATCTCTGCTACCAATGTCATCCAGACATCTTGATAAT	1389
DB	1090	ACTTTTGTGAGTGTGGTAGATCTCTGCTACCAATGTCATCCAGACATCTTGATAAT	1149
QY	1390	ACCTGTTGGAATCTCCTATGATCGCAATTAATCGAAGTCACTCAATGGTGAACACTGGG	1449
DB	1150	ACCTGTTGGAATCTCCTATGATCGCAATTAATCGAAGTCACTCAATGGTGAACACTGGG	1209
QY	1450	CAAACTAAGAGAATGAGAGCTGTGGAGTGTAGCAAGAGTGTATTAGAAATGTTACG	1509
DB	1210	CAAACTAAGAGAATGAGAGCTGTGGAGTGTAGCAAGAGTGTATTAGAAATGTTACG	1269
QY	1510	AAAAAACTATGTTGTGTCGAGTGGATTTTCACAGCCATTTCTTAAAGGAATATTT	1569
DB	1270	AAAAAACTATGTTGTGTCGAGTGGATTTTCACAGCCATTTCTTAAAGGAATATTT	1329
QY	1570	AGAAAGCCAAAGTCAGAAACCGGTGCTCTCTACTTTTCCCTGGAGCAAGCGTTGTACC	1629
DB	1330	AGAAAGCCAAAGTCAGAAACCGGTGCTCTCTACTTTTCCCTGGAGCAAGCGTTGTACC	1389
QY	1630	AGCTATCTTCTTCAAGACCCAGTGATGCTGCTGATGAAGGTGAGAGACAGTCCATTA	1689
DB	1390	AGCTATCTTCTTCAAGACCCAGTGATGCTGCTGATGAAGGTGAGAGACAGTCCATTA	1449
QY	1690	TGAGTCCAGAAATGCAACAGATGAATCCCTCAGAGGAGTGTGATTCGAAATCTGCTGA	1749
DB	1450	TGAGTCCAGAAATGCAACAGATGAATCCCTCAGAGGAGTGTGATTCGAAATCTGCTGA	1509
QY	1750	GCATATTTCTATTCATCTGTAGCAAGTCTGTGCAATTTATGTCACACACATTTGGCTTG	1809
DB	1510	GCATATTTCTATTCATCTGTAGCAAGTCTGTGCAATTTATGTCACACACATTTGGCTTG	1569
QY	1810	CCTGCTCTCTACAGACAGGAGGAAATGATCTCTCCCAATTTGGTGGAGACTTCTT	1869
DB	1570	CCTGCTCTCTACAGACAGGAGGAAATGATCTCTCCCAATTTGGTGGAGACTTCTT	1629
QY	1870	TGTGATGAAGAGGAGTCTGCTGATTTTGGACCTGGGGTCTCAGGAAATTCAGA	1929
DB	1630	TGTGATGAAGAGGAGTCTGCTGATTTTGGACCTGGGGTCTCAGGAAATTCAGA	1689
QY	1930	AGATGTAGTAATGTCATGCCATACAGCTGCTGGGAAATTTGTGCACAAATCACCCACTAG	1989
DB	1690	AGATGTAGTAATGTCATGCCATACAGCTGCTGGGAAATTTGTGCACAAATCACCCACTAG	1749
QY	1990	CAGGAAACGATGATTTTTTATCACCCCGACACAACTGTCCCATCAGTCTTCGAACTCAA	2049
DB	1750	CAGGAAACGATGATTTTTTATCACCCCGACACAACTGTCCCATCAGTCTTCGAACTCAA	1809
QY	2050	CTTCTACAGATGGGTACTTCTATCTTTATCATGAGGAGCCATCATAGCTTGCAGCCT	2109
DB	1810	CTTCTACAGATGGGTACTTCTATCTTTATCATGAGGAGCCATCATAGCTTGCAGCCT	1869
QY	2110	TTATGAGTCTTCTGAACAGAGGGGACTGGGGGGTCCCACTAGCACCCCACTAACTGAT	2169







RESULT 6  
AAL55475  
ID AAL55475 standard; DNA; 2487 BP.  
XX  
AC AAL55475;  
XX  
DT 22-MAY-2003 (first entry)  
XX  
DE GPAM related DNA sequence, SEQ ID No 1.  
XX  
KW Antidiabetic; nephrotropic; neuroprotective; ophthalmological; human;  
KW mitochondrial sn-glycerol-3-phosphate acyltransferase; GPAM;  
KW diabetic complication; retinopathy; neuropathy; enzyme; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 1..2487  
FT /\*tag= a  
ET /product= "GPAM-related protein"  
XX  
PN WO2003008590-A1.  
XX  
PD 30-JAN-2003.  
XX  
PF 16-JUL-2002; 2002WO-JP07189.  
XX  
PR 16-JUL-2001; 2001JP-0215337.  
XX  
PA (KISP) KISSEI PHARM CO LTD.  
XX  
PI Sakamoto S, Onota H, Sugano S, Nakamura Y;  
XX  
DR WPI; 2003-229583/22.  
DR P-PSDB; AAO27073.  
XX  
PT Human mitochondrial sn-glycerol-3-phosphate acyltransferase and  
PT antagonists for treatment and prevention of diabetic complications -  
XX  
PS Claim 4; Page 33-40; 56pp; Japanese.  
XX  
CC The invention relates to a novel protein having human mitochondrial sn-  
CC glycerol-3-phosphate acyltransferase (GPAM) activity. The novel protein  
CC with GPAM activity can be used in the prevention and treatment of  
CC diabetic complications, including retinopathy and neuropathy, by  
CC administration of antagonists to human GPAM. This polynucleotide sequence  
CC represents a DNA sequence encoding a protein relating to the GPAM  
XX activity protein of the invention.  
XX  
SQ Sequence 2487 BP; 704 A; 549 C; 571 G; 663 T; 0 other;

Query Match 82.7%; Score 2483.8; DB 25; Length 2487;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2485; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 341 ATGGATGAATCTGCACCTGACCTGGTACATAGATGTTCTTATCTGCCACATTCATCA 400  
Db 1 ATGGATGAATCTGCACCTGACCTGGTACATAGATGTTCTTATCTGCCACATTCATCA 60  
Qy 401 GAATACAGTGTGGTTCGATGTAAGCACACAAGTGAAGGAATGGGTGAGTGGCTTTAGA 460  
Db 61 GAATACAGTGTGGTTCGATGTAAGCACACAAGTGAAGGAATGGGTGAGTGGCTTTAGA 120  
Qy 461 CCCACCGTCTTCAGATCTGCACTTTAAATGGAAGAAAGCCTTAATGAGTCGGAAGAGG 520  
Db 121 CCCACCGTCTTCAGATCTGCACTTTAAATGGAAGAAAGCCTTAATGAGTCGGAAGAGG 180  
Qy 521 CCATTGTTGGAAGATGTTGTTACTCTCGACTCCCGAGCTGGGACAAATTTTCAAC 580  
Db 181 CCATTGTTGGAAGATGTTGTTACTCTCGACTCCCGAGCTGGGACAAATTTTCAAC 240

Qy 581 CCCAGTATCCCGTCTTTGGGTTTCGGAATGTTATTTATATCAATGAACATCACACAAGA 640  
Db 241 CCCAGTATCCCGTCTTTGGGTTTCGGAATGTTATTTATATCAATGAACATCACACAAGA 300  
Qy 641 CACCGGATGGCTTGCAGACGCTTTCTTACGTTCTTTTATTTTATTTCAAGACGAGATGTG 700  
Db 301 CACCGGATGGCTTGCAGACGCTTTCTTACGTTCTTTTATTTTATTTCAAGACGAGATGTG 360  
Qy 701 CATTAAGGGCATGTTTGCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 760  
Db 361 CATTAAGGGCATGTTTGCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 420  
Qy 761 GAGCAATTCGAGAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCACGACGAATCA 820  
Db 421 GAGCAATTCGAGAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCACGACGAATCA 480  
Qy 821 AAAGCCGTTAAACAAGTGAAGAAAGAAAGCTTAAAGGATTTCTTCAAGAAATGGTTGCCACT 880  
Db 481 AAAGCCGTTAAACAAGTGAAGAAAGAAAGCTTAAAGGATTTCTTCAAGAAATGGTTGCCACT 540  
Qy 881 GTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGCTTAAACCTGTTCAACAGCTTC 940  
Db 541 GTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGCTTAAACCTGTTCAACAGCTTC 600  
Qy 941 TTTTGAACATTCAAATTCACAAAGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACG 1000  
Db 601 TTTTGAACATTCAAATTCACAAAGTCAACTTGAGATGGTTAAAGCTGCAACTGAGACG 660  
Qy 1001 AATTGCGGCTTCTGTTTCTACAGTTTCATAGATCCCATATTTGACTATCTGCTCACT 1060  
Db 661 AATTGCGGCTTCTGTTTCTACCAAGTTCATAGATCCCATATTTGACTATCTGCTCACT 720  
Qy 1061 TTCAATCTCTTCTGCCATAACATCAAGACCATACATTTGCTTCCAGCAATAATCTCAAC 1120  
Db 721 TTCAATCTCTTCTGCCATAACATCAAGACCATACATTTGCTTCCAGCAATAATCTCAAC 780  
Qy 1121 ATCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTTCATACGACGAAGGCTC 1180  
Db 781 ATCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTTCATACGACGAAGGCTC 840  
Qy 1181 GATGAACACACAGATGACGGAAGATGTTCTTATAGAGCTTTGCTTCCATGGGCATATA 1240  
Db 841 GATGAACACACAGATGACGGAAGATGTTCTTATAGAGCTTTGCTTCCATGGGCATATA 900  
Qy 1241 GTTGAATTTACTTGCACAGCAGCAATTTCTGGAGATCTTCTGGAGCAGCAGCTTCAGG 1300  
Db 901 GTTGAATTTACTTGCACAGCAGCAATTTCTGGAGATCTTCTGGAGCAGCAGCTTCAGG 960  
Qy 1301 AGTGGAAAAACCTCTTGTGCTCGGGCAGGACTTTTGTGAGTGTGGTAGATACCTGTCT 1360  
Db 961 AGTGGAAAAACCTCTTGTGCTCGGGCAGGACTTTTGTGAGTGTGGTAGATACCTGTCT 1020  
Qy 1361 ACCAATGTCTATCCAGACATCTTTGATTAATACCTGTTTGGAAATCTCTATGATCGCATATC 1420  
Db 1021 ACCAATGTCTATCCAGACATCTTTGATTAATACCTGTTTGGAAATCTCTATGATCGCATATC 1080  
Qy 1421 GAAGGTCACTACATGCTGGAACCTGGGCAACCTTAAGAAATGAGAGCTGTGAGT 1480  
Db 1081 GAAGGTCACTACATGCTGGAACCTGGGCAACCTTAAGAAATGAGAGCTGTGAGT 1140  
Qy 1481 GTAGCAAGAGTGTATTAGAAATGTACGAAAAAATCTATGTTGTGTCGAGTGGATTTT 1540  
Db 1141 GTAGCAAGAGTGTATTAGAAATGTACGAAAAAATCTATGTTGTGTCGAGTGGATTTT 1200  
Qy 1541 GCACAGCCATTTTCTTAAAGGAATATTAGAAAGCCAAAGTTCAGAAACCGGTGTCTGCT 1600  
Db 1201 GCACAGCCATTTTCTTAAAGGAATATTAGAAAGCCAAAGTTCAGAAACCGGTGTCTGCT 1260  
Qy 1601 CTACTTTTCCCTGGAGCAGGTTGTTACAGCTATACCTCTTCAAGACCCAGTGTGCT 1660  
Db 1261 CTACTTTTCCCTGGAGCAGGTTGTTTACAGCTATACCTCTTCAAGACCCAGTGTGCT 1320  
Qy 1661 GCTGATGAAGGTAGACACGTCCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1720

Db	1321	GCTGATGAAGGTTAGAGACAGTCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA	1380
Qy	1721	CGAAGGAGGTTGATTGCAAACTCTGGCTGAGCATATCTTAATCACTGCTAGCAAGTCTGT	1780
Db	1381	CGAAGGAGGTTGATTGCAAACTCTGGCTGAGCATATCTTAATCACTGCTAGCAAGTCTGT	1440
Qy	1781	GCCATTATGTCACACACATTTGTGGCTTGCTGCTCTTACAGACACAGCAGGGAATT	1840
Db	1441	GCCATTATGTCACACACATTTGTGGCTTGCTGCTCTTACAGACACAGCAGGGAATT	1500
Qy	1841	GATCTCTCCCATTTGGTTCGAAGACTTCTTTGTGATGAAGAGGAGTCTTGGCTCTGTGAT	1900
Db	1501	GATCTCTCCCATTTGGTTCGAAGACTTCTTTGTGATGAAGAGGAGTCTTGGCTCTGTGAT	1560
Qy	1901	TTTGACCTGGGTTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGCTGCTG	1960
Db	1561	TTTGACCTGGGTTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGCTGCTG	1620
Qy	1961	GGAAATTTGTGTCAAACTCACCCACATAGCAGGAACGATGAGTTTTTATATCACCCCCAGC	2020
Db	1621	GGAAATTTGTGTCAAACTCACCCACATAGCAGGAACGATGAGTTTTTATATCACCCCCAGC	1680
Qy	2021	ACAACCTGTCATCAGTCTTCGAACCTCAACTTTACAGCAATGGGGTACTTCAATGCTTT	2080
Db	1681	ACAACCTGTCATCAGTCTTCGAACCTCAACTTTACAGCAATGGGGTACTTCAATGCTTT	1740
Qy	2081	ATCATGAGGCCATCATAGCTTGAGCCTTTATGACGTTCTGAACAGAGGGGACGTGGG	2140
Db	1741	ATCATGAGGCCATCATAGCTTGAGCCTTTATGACGTTCTGAACAGAGGGGACGTGGG	1800
Qy	2141	GGTCCCACTAGCACCCACCTAACTGATCAGCCAGGAGCAGCTGGTGCAGAAAGCGGCC	2200
Db	1801	GGTCCCACTAGCACCCACCTAACTGATCAGCCAGGAGCAGCTGGTGCAGAAAGCGGCC	1860
Qy	2201	AGCCTGTGCTACCTTCTCTCAATGAAGGCAATCTCACTGCCTTGCCAGACATTTTAC	2260
Db	1861	AGCCTGTGCTACCTTCTCTCAATGAAGGCAATCTCACTGCCTTGCCAGACATTTTAC	1920
Qy	2261	CNAGTCTGCCATGAACAGTAGGAAAGTTATCCAGTATGGCATCTTACAGTGGCAGAG	2320
Db	1921	CNAGTCTGCCATGAACAGTAGGAAAGTTATCCAGTATGGCATCTTACAGTGGCAGAG	1980
Qy	2321	CACGATGACAGGAAGATATCAGTCTTAGTCTTGCTGAGCAGCAGTGGGACAGAAAGCTT	2380
Db	1981	CACGATGACAGGAAGATATCAGTCTTAGTCTTGCTGAGCAGCAGTGGGACAGAAAGCTT	2040
Qy	2381	CCAGAACCTTTGCTCTGGAGAAGTGATGAAGAAGATGAAGACAGTACTTTGGGGAGGAA	2440
Db	2041	CCTGAACTTTGCTCTGGAGAAGTGATGAAGAAGATGAAGACAGTACTTTGGGGAGGAA	2100
Qy	2441	CAGCGAGATGCTACTGTAAGGTTAGCCATCCAGGAGCAGCAGCAGTATATCACCTTC	2500
Db	2101	CAGCGAGATGCTACTGTAAGGTTAGCCATCCAGGAGCAGCAGCAGTATATCACCTTC	2160
Qy	2501	TTACAGAGACTCTCTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCATCTTTGTTTAC	2560
Db	2161	TTACAGAGACTCTCTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCATCTTTGTTTAC	2220
Qy	2561	AACTTCAGTGGTCTGTTTCCAGAACCTGAGTATCTGAAAAGTTGCAAAATCTCTAATA	2620
Db	2221	AACTTCAGTGGTCTGTTTCCAGAACCTGAGTATCTGAAAAGTTGCAAAATCTCTAATA	2280
Qy	2621	ACCAGAACAGAAAAGTATGCTGAGAGTGCCACATATTTGCTTGTGAAG	2680
Db	2281	ACCAGAACAGAAAAGTATGCTGAGAGTGCCACATATTTGCTTGTGAAG	2340
Qy	2681	AATCCTGTGAAAATGTTTAAAGTATTTGGGTTTTTCAAGGAGACCAACAAAAGAGAGTG	2740
Db	2341	AATCCTGTGAAAATGTTTAAAGTATTTGGGTTTTTCAAGGAGACCAACAAAAGAGAGTG	2400
Qy	2741	TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCCAACCGCAAAAATCTTCTAGAA	2800

Db	2401	TCGTGTTTGAACCTAGCAGCACCTTTTCTACCTCAATGCAACGCAAAAACCTTCTAGAA	2460
Qy	2801	TATATCTGAGTTTTGTGCTGTAG	2827
Db	2461	TATATCTGAGTTTTGTGCTGTAG	2487
RESULT 7			
ABZ12025			
ID	ABZ12025	standard; cDNA; 3273 BP.	
XX	AC	ABZ12025;	
XX	XX	20-JAN-2003 (first entry)	
XX	XX	Human polynucleotide SEQ ID NO 907.	
XX	XX	Human; genome mapping; gene therapy; food supplement; virus; fungus;	
KW	KW	cell-proliferative disorder; neurodegenerative disease; bacterial;	
KW	KW	Parkinson's disease; Alzheimer's disease; autoimmune disease;	
KW	KW	multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;	
KW	KW	arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;	
KW	KW	antiparkinsonian; antidiabetic; immunosuppressive; dermatological;	
KW	KW	haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;	
KW	KW	antiarthritic; gene; ss.	
XX	XX		
OS	OS	Homo sapiens.	
XX	XX		
XX	XX	WO200270539-A2.	
XX	XX	12-SEP-2002.	
XX	XX	05-MAR-2002; 2002WO-US05095.	
XX	XX	05-MAR-2001; 2001US-0799451.	
XX	XX	(HYSE-) HYSEQ INC.	
XX	XX		
PI	PI	Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;	
PI	PI	Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;	
PI	PI	Wehrman T, Wang J, Wang D, Drmanac RT;	
XX	XX	WPI; 2002-759812/82.	
DR	DR	P-PSDB; ABP69808.	
XX	XX		
PT	PT	New polynucleotides comprising sequences assembled from expressed	
PT	PT	sequence tags (ESTs), useful for treating cell-proliferative,	
PT	PT	neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or	
PT	PT	platelet or coagulation disorders -	
XX	XX		
PS	PS	Claim 1; SEQ ID NO 907; 1012pp + Sequence Listing; English.	
XX	XX		
CC	CC	The invention relates to an isolated polynucleotide (I) comprising a	
CC	CC	nucleotide sequence selected from any of 948 sequences	
CC	CC	(ABZ11119-ABZ12066) or their mature protein coding portion, active domain	
CC	CC	coding protein or complementary sequences. The polynucleotides are useful	
CC	CC	for identifying expressed genes or for physical mapping of human genome.	
CC	CC	The encoded polypeptides (ABP68902-ABP69849) are useful as molecular	
CC	CC	weight markers, as a food supplement, for generating antibodies, in	
CC	CC	medical imaging, screening and diagnostic assays and for treating	
CC	CC	cell-proliferative disorders (cancer), neurodegenerative diseases	
CC	CC	(Parkinson's or Alzheimer's disease), autoimmune diseases (multiple	
CC	CC	sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid	
CC	CC	disorders, platelet or coagulation disorders, wound burns, incision,	
CC	CC	ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,	
CC	CC	parasitic), arthritis, etc.	
CC	CC	Note: The sequence data for this patent did not form part of the printed	
CC	CC	specification, but was obtained in electronic format directly from WIPO	
CC	CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX	XX		
SQ	XX	Sequence 3273 BP; 915 A; 704 C; 733 G; 921 T; 0 other;	

Best Local Similarity 99.6%; Pred. No. 0; Matches 2359; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY	636	CAAGACACCGCGATGGCTTGGCAAGACGCGCTTCTTACGTTCTTTTATTCAAGACGAG	695	1193	ATGCTGCTGATGAAGGTAGAGACACGTCCTCAATTAATGATCCAGAAATCAACAGATGAAT	1252
DB	173	CCAGACACCGCGATGGCTTGGCAAGACGCGCTTCTTACGTTCTTTTATTCAAGACGAG	232	1716	CCCTACGAAGGAGGTTCGATTCGCAATCTGGCTGAGCATATTTCTATTCACTCTAGCAAGT	1775
QY	696	ATGTGCATAAAGGCGATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAG	755	1253	CCCTACGAAGGAGGTTCGATTCGCAATCTGGCTGAGCATATTTCTATTCACTCTAGCAAGT	1312
DB	233	ATGTGCATAAAGGCGATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAG	292	1776	CCTGTGCCATTATGTCACACACATTTGGCTTGCCTCTCTCTACAGACACAGGCGAG	1835
QY	756	TACAGAGGCATTCAGAGAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCACAGCAGC	815	1313	CCTGTGCCATTATGTCACACACATTTGGCTTGCCTCTCTCTACAGACACAGGCGAG	1372
DB	293	TACAGAGGCATTCAGAGAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCACAGCAGC	352	1836	GAATGTATCTCTCCACATTTGGTCCAGACCTTCTTTGTGATGAAGAGGAAGTCTCTGGCTC	1895
QY	816	AATCAAAAGCCGTTAAACAAAGTGAAGAAAGAGTCTTCAAGAAATGGTTG	875	1373	GAATGTATCTCTCCACATTTGGTCCAGACCTTCTTTGTGATGAAGAGGAAGTCTCTGGCTC	1432
DB	353	AATCAAAAGCCGTTAAACAAAGTGAAGAAAGAGTCTTCAAGAAATGGTTG	412	1896	GTGATTTTGACCTCGCGGTTCTCAGCAATTCAGAAATTCAGAAAGATGTAGTAATGATGCCATACAGC	1955
QY	876	CCACTGTCTACCCGCAATGATCAGACTGACTGGGTGGTGTCTGCTAAAACCTGTTCAACA	935	1433	GTGATTTTGACCTCGCGGTTCTCAGCAATTCAGAAATTCAGAAAGATGTAGTAATGATGCCATACAGC	1492
DB	413	CCACTGTCTACCCGCAATGATCAGACTGACTGGGTGGTGTCTGCTAAAACCTGTTCAACA	472	1956	TGCTGGGAAATTTGTTCACAAATCAACCCACACTAGCAGGAACGATGAGTTTTTATCACCC	2015
QY	936	GCTTCTTTTGGACATTCAAATTCACAAAGGTCAACTTGAGATGTTAAAGCTGCAACTG	995	1493	TGCTGGGAAATTTGTTCACAAATCAACCCACACTAGCAGGAACGATGAGTTTTTATCACCC	1552
DB	473	GCTTCTTTTGGACATTCAAATTCACAAAGGTCAACTTGAGATGTTAAAGCTGCAACTG	532	2016	CCAGACAACTGTCTCCCATCAGTCTTTCGAACTCAACCTCTACAGCAATGGGGTACTTCATG	2075
QY	996	AGACAAATTTGCGGCTTCTGTTTCTACAGATTCATAGATCCCATATGACTATCTGCTGC	1055	1553	CCAGACAACTGTCTCCCATCAGTCTTTCGAACTCAACCTCTACAGCAATGGGGTACTTCATG	1612
DB	533	AGACAAATTTGCGGCTTCTGTTTCTACAGATTCATAGATCCCATATGACTATCTGCTGC	592	2076	TCCTTTATCATGGAGGCCATCATAGCTTTCGAGCCCTTTATGCAAGTCTGAAACAAAGAGGGAC	2135
QY	1056	TCACATTTTCAATCTCTTCTGCCATAACATCAAGACACCATACATTTGCTTCAGGCAATATC	1115	1613	TCCTTTATCATGGAGGCCATCATAGCTTTCGAGCCCTTTATGCAAGTCTGAAACAAAGAGGGAC	1672
DB	593	TCACATTTTCAATCTCTTCTGCCATAACATCAAGACACCATACATTTGCTTCAGGCAATATC	652	2136	TGGGGGGTCCCATPAGCACCCACCTAACCCTGATCAGCAGGAGCAGCTGGTGGGGAAGG	2195
QY	1116	TCAACATCCCAATCTTCTAGTACCTTGATCCATAAGCTTGGGGCTTCTTCATACGACGAA	1175	1673	TGGGGGGTCCCATPAGCACCCACCTAACCCTGATCAGCAGGAGCAGCTGGTGGGGAAGG	1732
DB	653	TCAACATCCCAATCTTCTAGTACCTTGATCCATAAGCTTGGGGCTTCTTCATACGACGAA	712	2196	CGGCAGCGCTGTGCTACCTTCTTCCAAATGAAGGCAACATCTCACTGCTGCTGCCAGACAT	2255
QY	1176	GSCTCGATGAACACACGATGACGAGGAGGATGTTCTCTATAGAGCTTTGCTCCATGGGC	1235	1733	CGGCAGCGCTGTGCTACCTTCTTCCAAATGAAGGCAACATCTCACTGCTGCTGCCAGACAT	1792
DB	713	GSCTCGATGAACACACGATGACGAGGAGGATGTTCTCTATAGAGCTTTGCTCCATGGGC	772	2256	TTTACCAGTCTGCCATGAACACGATGAGGAAAGTTTATCCAGTATGGCATCTTACAGTGG	2315
QY	1236	ATATAGTTGAATTTACTTCGACAGCAGCAATTTCTGGAGATCTTCCTGGAAGGCAACGCTT	1295	1793	TTTACCAGTCTGCCATGAACACGATGAGGAAAGTTTATCCAGTATGGCATCTTACAGTGG	1852
DB	773	ATATAGTTGAATTTACTTCGACAGCAGCAATTTCTGGAGATCTTCTGGAAGGCAACGCTT	832	2316	CAGACACGATGACCCAGGAAAGATATCAGTCTTCTAGTCTTGTGCTGAGCAGCAGTGGGCAAGA	2375
QY	1296	CTAGAGTGGAAAAACCTCTTGTGCTCGGCGAGGACTTTTGTGCTAGTGTGTAGATCTC	1355	1853	CAGACACGATGACCCAGGAAAGATATCAGTCTTCTAGTCTTGTGCTGAGCAGCAGTGGGCAAGA	1912
DB	833	CTAGAGTGGAAAAACCTCTTGTGCTCGGCGAGGACTTTTGTGCTAGTGTGTAGATCTC	892	2376	AGCTTCCAGAACCTTTGTCTTGGAGAAAGTGAAGAAAGATGAAGACAGTGCATTTGGGG	2435
QY	1356	TGCTTACCAATGTCTCCAGACATCTTGATAATACCTTTGGATCTCCTATGATCGCA	1415	1913	AGCTTCCAGAACCTTTGTCTTGGAGAAAGTGAAGAAAGATGAAGACAGTGCATTTGGGG	1972
DB	893	TGCTTACCAATGTCTCCAGACATCTTGATAATACCTTTGGATCTCCTATGATCGCA	952	2436	AGGAACAGCGAGATTGCTTACCTGAAGGTGAGCCCAATCCAGGAGCAGCAGTTCATCA	2495
QY	1416	TTATCGAAGGTCACTACAAATGTGAAACAACTGGGCAAACTTGAAGAAATGAGAGCCTGT	1475	1973	AGGAACAGCGAGATTGCTTACCTGAAGGTGAGCCCAATCCAGGAGCAGCAGTTCATCA	2032
DB	953	TTATCGAAGGTCACTACAAATGTGAAACAACTGGGCAAACTTGAAGAAATGAGAGCCTGT	1012	2496	CCTTCTTACAGAGACTCCTTTGGGCTTGTGCTGAGGCTTACAGCTCTGCTGCCATCTTTG	2555
QY	1476	GGAGTGTAGCAAGAGGTGTTATAGAATGTTACGAAAAAATATGTTGTTGTGCGAGTGG	1535	2033	CCTTCTTACAGAGACTCCTTTGGGCTTGTGCTGAGGCTTACAGCTCTGCTGCCATCTTTG	2092
DB	1013	GGAGTGTAGCAAGAGGTGTTATAGAATGTTACGAAAAAATATGTTGTTGTGCGAGTGG	1072	2556	TTCAAACTTCAGTGTCTGTTCCAGAACTGAGTATCTGCAAAAGTGTGCAAAATACC	2615
QY	1536	ATTTTGCACAGCCATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCCAGAAACCGGTGT	1595	2093	TTCAAACTTCAGTGTCTGTTCCAGAACTGAGTATCTGCAAAAGTGTGCAAAATACC	2152
DB	1073	ATTTTGCACAGCCATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCCAGAAACCGGTGT	1132	2616	TAATAACAGCAAGAGAAATGTTTCAGTATATGCTGAGAGTGCACATATTTGTCTTTG	2675
QY	1596	CTGCTCTACTTTCCTCGAGCAAGCTTGTGTACAGCTATATCTTCCCTCAAGACCCAGTG	1655	2153	TAATAACAGCAAGAGAAATGTTTCAGTATATGCTGAGAGTGCACATATTTGTCTTTG	2212
DB	1133	CTGCTCTACTTTCCTCGAGCAAGCTTGTGTACAGCTATATCTTCCCTCAAGACCCAGTG	1192	2676	TGAAGAATGCTGTGAAATGTTTTAAGGATATTTGGGGTTTTTCAAGAGACCCAAACAGA	2735
QY	1656	ATGCTGCTGATGAAGGTAGACACGCTCCATTAATGAGTCCAGAAATGCAACAGATGAAT	1715	2213	TGAAGAATGCTGTGAAATGTTTTAAGGATATTTGGGGTTTTTCAAGAGACCCAAACAGA	2272
DB				2736	GAGTGTCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAATGCAACCGACAAACCTTC	2795
				2273	GAGTGTCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAATGCAACCGACAAACCTTC	2332



Qy	1039	TATTGACGTATCTGCTGCTCACTTTTCATTCTCTTCTGCGCATAACTCAAGGACCATACAT	1098
Db	858	CATCGACTACCTGCTGCTCACTTTCATCTCTTCTGCGCAACAACATCAAGCTCCATACAT	917
Qy	1099	TGCTTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTGATCCATAAGCTTTGGGG	1158
Db	918	CGCTCGGGCACAACCTCAACATCCCACTTCCTCAGTACCTTGATTCACAAGCTTGGGG	977
Qy	1159	CTTCTTCATACGACGAAGGCTCGATGAAACACCAAGATGGACGGAAGATGTTCTCTATAG	1218
Db	978	CTTTTTCATAGACGGAGGCTTGACGAAACTCCAGATGGACGGAAGACATTCGTACAG	1037
Qy	1219	AGCTTTCCTCATCGGCATATAGTTGAATTTTCTGACACGACGCAATTCCTGGAGATCTT	1278
Db	1038	AGCGTCTCTCATGGGCATATAGTTGAACTCTCTCGACAGCAGGATTCCTGAGATCTT	1097
Qy	1279	CCTGGAAGGCACAGCTTCTAGGAGTGAAAAACCTCTGTGCTCGGCGCAGCACTTTTGTC	1338
Db	1098	CCTGGAAGGCACCCGCTCCGCGAGTGGCAAGACCTCTGTGCCGGCGCGGCTCCTGTC	1157
Qy	1339	AGTTGTGGTAGATCTCTGTCTTACCAATGTCACTCCAGACATCTTGATAATACTCTGTTGG	1398
Db	1158	AGTGSTAGTGATATCTGTGATCCAAACCATCCCTGATCATCTGTGTCATCCCTGTGGG	1217
Qy	1399	AATCTCTATGATCGCATATCGAAGTCACTACAATGGTGAACAACCTGGGCAACCTTAA	1458
Db	1218	CATCTGTATGATCGGATAATCGAAGTCACTACAATGGTGAACAGTGGGCGAAGGCCAA	1277
Qy	1459	GAAGATGAGAGCCGTGTGGAGTGTAGCAAGAGGTGTTATTAGAAATGTTCGAAAAACTA	1518
Db	1278	GAAGATGAAGTCTCTGGAGTGTGGCAAGGCGTTATCAGAATGCTCGGAAAAACTA	1337
Qy	1519	TGTTGTGTCGAGTGAATTTTGACAGCAATTTTCCCTTAAAGAAATTTTGAAGAGCCA	1578
Db	1338	CGGCTATGTCCGAGTGAATTTGACAGCAATTTTCTTCAAGGAAATTTTGAAGAGCCA	1397
Qy	1579	AAGTCAGAAAACGGTGTCTGCTACTTTCCCTGGAGCAAGCGTTGTACCGACTACT	1638
Db	1398	AAGTCAGAAAACCTGTATCTGCTCCCTCTCTTTGGAGCAAGCACTGTTCAGCAATCT	1457
Qy	1639	TCCTTCAAGACCCAGTGATGCTGTGATGAAGGTAGAGACACGCTCCAATTAATGAGTCCAG	1698
Db	1458	TCCTTCAAGACCTGTATGCTGCTGCGCAACATGAAGATGTCCAGTAATGAGTCGAG	1517
Qy	1699	AAATGCAACAGATGAATCCCTTAAGAGAGGTGATTTGCAAAATCTGGCTGAGCATATCT	1758
Db	1518	AAACGCGGCAGACGAAGCTTCCGAAGAGGCTGATCGCAAACTGGCGGAGCACATCT	1577
Qy	1759	ATTCACTGCTAGCAAGTCTGTGCAATATGTTCACACACATTTGTGGCTTGCGTCTCT	1818
Db	1578	CTTCACCGCAGCAAGTCTGTGCGTATCATGTCCACCACATTTGTGGCCCTGCTGTCTCT	1637
Qy	1819	CTACAGACACAGCAGGGAATGATCTCTCCACATTTGGTGGAAAGACTTCTTTTGTGATGA	1878
Db	1638	CTACAGACACAGCAGGGAATCCACCTCTCCACGCTGTGGAAGACTTCTTTTGTGATGA	1697
Qy	1879	AGAGGAATCTTGCTGCTGTGATTTTGACCTGGGGTCTCAGGAAATTCAGAAATGTAGT	1938
Db	1698	GGAGGAATCTTACGCTCGGAATTTTGACCTGGGCTCTCCGGGAAATTCAGAAATGTAGT	1757
Qy	1939	AATGCATGCCATACAGCTGCTGGGAAATTTGTTCACAATCACCACACTAGCAGGAACGA	1998
Db	1758	CATGATGCTATTACGCTTCTGGGGAATGTGTCACAATCACCCCACTAGCAGGAAGGA	1817
Qy	1999	TGAGTTTTTATCACCCCCGACCAACTGTGCCATCATGTTCTCGAACTCAACTTCTACAG	2058
Db	1818	TGAATCTTTTATTACTCCAGCAAACTGTCCCGTCTCTTGGAACTCAACTTCTACAG	1877
Qy	2059	CAATGGGGTACTCATGCTTTTATCATGGAGGCATCATAGCTTGAGCCTTTATGCAGT	2118
Db	1878	CAATGGGGTACTTCATGCTTTTATCATGGAAGGCATCATAGCTTGAGCATTTATGCAGT	1937

QY	2119	TCTGAACAAGAGGGGACTGGGGGTCCTACTAGCACCCACCTTAACCTGATCAGCAGGA	2178
DB	1938	CCAGAAATAAGAGGGGTTCCGGAGGGCTGTCGGGAGGCCTTGGCAACTCTGATCAGCCAGGA	1997
QY	2179	GCAGCTGGTCGGNAAGCGCCAGCCTGTGCTACTCTTCTCTCAANTGAAGGACCAATCTC	2238
DB	1998	GCAGCTGGTCGGAGAGCGCCAGCCTGTGCTACTCTCTCTAAATGAAGGTACCAATTC	2057
QY	2239	ACTGCCCTTGCACAGACATTTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTA	2298
DB	2058	TCTGCCCTGCCAGACATTTTACCAAGTTTGTCAAGAGACAGTAGGAAAGTTTATCCAGTA	2117
QY	2299	TGGCAATCTTACAGTGGCAGAGCACGATACCAAGGAAGATATACGTCCTTAGTCTTGCTGA	2358
DB	2118	CGGAAATCTCACAGTGGCAGAGCAAGATACCAGGAAGATGTCAGTCTCGGCCCTTCGAGA	2177
QY	2359	GCAGCAGTGGGACAAAGNAGCTTCCAGAACCTTTGTCTTGGAGNAGTGTAGAGAAATGA	2418
DB	2178	GCAGCAGTGGAAACAAGAAGCTTCCGAGCGCTCTGAACTCGGAGAAGTGACGAAGAAATGA	2237
QY	2419	AGACAGTGAATTTTGGGAGGAAACAGCGAGATTCCTACTGAAAGTCAGCCAAATCCAAAGGA	2478
DB	2238	GGACAGTGAATTTTGGTAGAGAGCAGGTGATTCCTACTGAAGGTGAGCCAGGCCAAGGA	2297
QY	2479	GCACCAGCAGTTTATCACCTTTTACAGAGACTCCTTTGGGCGCTTTGCTGGAGCCTTACAG	2538
DB	2298	GCACCAGCAATTCATCACCTTTCTGAGAGGCCTTCTGGGCGCCCTGCTAGAGCCTTACAG	2357
QY	2539	CTCTGCTGCCATCTTTGTTCCAACTTCAGTGTCTCTGTTTCCAGAAACCTTAGTATCTGCA	2598
DB	2358	CTCTGCTGCCATCTTTGTTCCACACCTTCCGCGCCCAAGTCCCGAGTCTGAGTACCTGCA	2417
QY	2599	AAAGTTGCCACAAATACCTAATACCAGACAGAAAGAAATGTTGTCAGTATATGCTCAGAG	2658
DB	2418	GAAGCTGCACAGGTACTCTCTCCACCAGGACGGAGAGGAACGTCGCGGTGTACGCTGAGAG	2477
QY	2659	TGCCACATATGTCCTTGTGAAGAAATGCTGTGAAAATGTTTAAAGGATATTTGGGTTTTCAA	2718
DB	2478	TGCCACATACTGTCCTTGTGAGAAATGCTGTGAAAATGTTTAAAGCATATCGGGGTTTTCAA	2537
QY	2719	GGAGACCAAAACAAGAGAGTGTCTGCTTTTAGAACTGAGCAGCACATTTTCTACCTCAATG	2778
DB	2538	AGAGACCAACGAGACGACGCTCTGCTTAGAACTGAGCACCATTTCTCTACCTCAGG	2597
QY	2779	CAACCGCAAAAACTCTTAGAATATATCTCGAGTTTTTGGTGCTGTAG	2827
DB	2598	CAGCCGCGCAAGCTCTCTGAAATACATCTCGAGCTTCTGAGCTTCGTGGTGCTGTAG	2646

PRECIT.T 9

RESULTS 9  
ABR09415

ABA08415  
ID: ABA08415 standard: cDNA: 469 BP.

ID  
 YY  
 ABA08415

XX  
AC  
ABA09415:

AC  
YY  
ABA08415;

DT 11-JAN-2002 (first entry)

11-JAN-20  
DY  
YYXX  
DE Human sn-glycerol-3-P acyltransferase homologue cDNA. SEO ID NO:191.

DE human vit-glyceral-3-P acyltransferase homologous cDNA, SEQ ID NO:151: XX

KW Human; cytokine; cell proliferation; cell differentiation; growth factor;

KW haematopoiesis regulation; tissue growth; immunomodulator; activin; immunomodulation; growth

KW inhibit; chemotaxis; chemokinesis; thrombolysis; oncogenesis;

KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;

KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW chronic inflammatory condition; malformation; not in mother

KW chronic inflammatory condition; proliferative retinopathy;  
KW atherosclerosis; coronary heart disease; arterial ischaemia.

KW bone disorder: osteoporosis: vascular growth disorder:  
KW atherosclerosis; coronary heart disease; arterial ischaemia;  
KW bone disorder: osteoporosis: vascular growth disorder:

bone disorder, osteoporosis, vascular growth disorder;  
tissue regeneration: wound healing: immune disorder:

KW cell culture; drug screening; gene therapy; antiinflammatory;

KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;

KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacteri

antifungal; vulnerary; antiulcer; ss.





CC Also included are: (1) a transformed cell having a nucleic acid  
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-  
CC translated sequence that functions in the cell to cause termination of  
CC transcription and addition of polyadenylated ribonucleotides to a 3' end  
CC of the mRNA molecule; and (2) determining a level or pattern of a  
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker  
CC nucleic acid comprising any of the 15112 nucleic acid sequences or its  
CC complement or fragment) with a complementary nucleic acid molecule  
CC obtained from the bovine cell or tissue, where hybridisation between the  
CC marker nucleic acid and the complementary nucleic acid permits the  
CC detection of the molecule; and (b) detecting the level or pattern of the  
CC complementary nucleic acid, where the detection of the complementary  
CC nucleic acid is predictive of the level or pattern of the molecule.  
CC The LMFD nucleic acid is used for determining a level or pattern  
CC of a molecule in a bovine cell or tissue. It is useful for genome  
CC mapping, gene identification and analysis, cattle breeding, preparation  
CC of constructs for use in cattle gene expression, or for genetically  
CC improving cattle. The present sequence is one of the 15112 bovine  
CC LMFD EST (expressed sequence tag) nucleic acids.  
CC Note: The present sequence was not shown in the specification but  
CC was obtained in electronic format from the USPTO web site:  
CC seqdata.uspto.gov/sequence.html?DocID=20020137139.  
XX  
SQ Sequence 429 BP; 113 A; 106 C; 94 G; 116 T; 0 other;

Query Match 11.4%; Score 342.2; DB 25; Length 429;  
Best Local Similarity 90.2%; Pred. No. 2e-95;  
Matches 378; Conservative 0; Mismatches 38; Indels 3; Gaps 1;  
QY 1670 GGTAGACACAGTCCATTAAAGTCCAGAAATGCAACAGATGAATCCTAGAGGAGG 1729  
DB 14 GATACAGACAGTCCATTAAAGTCCAGAAATGCAACAGATGAATCCTAGAGGAGG 70  
QY 1730 TTGATTGCAAAATCTGGCTGAGCATATCTATTCACTGCTAGCAAGTCTGTGCCATTATG 1789  
DB 71 CTGATTGACATCTGGCTGAGCACAATCTTCTACTGCTAGCAAGTCTGTGCCATTATG 130  
QY 1790 TCCACACACATTTGGTGGTGGCTGCTCTCTACAGACACAGGAGGAAATGATCTCTCC 1849  
DB 131 TCCAGGCACATCGTACCTGTCTGCTCTCTACAGACACAGGAGGAAATGCTCTCC 190  
QY 1850 ACATTTGTCGAGACTCTTTTGATGAAGAGGAGTCTGGCTGCTGATTTGACCTG 1909  
DB 191 ACATTTGTCGAGACTCTTTTGATGAAGAGGAGTCTGGCTGCTGATTTGACCTG 250  
QY 1910 GGGTCTCTCAGGAAATTCAGAGATCTAGTAATGATGCTATGATGCTGCTGGAAATGT 1969  
DB 251 GGTCTCTCAGGAAATTCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 310  
QY 1970 GTCAATCACCACACTAGCAGGAGATGATGATGATGATGATGATGATGATGATGATGAT 2029  
DB 311 ATCAATCACCACACTAGCAGGAGATGATGATGATGATGATGATGATGATGATGATGAT 370  
QY 2030 CCATCAGTCTTGAAGTCACTCACTTCTACAGCAATGGGGTACTTCTATCATGGA 2088  
DB 371 GCATCAGTCTTGAAGTCACTCACTTCTACAGCAATGGGGTACTTCTATCATGGA 429

## RESULT 11

ABX43656  
XX ID ABX43656 standard; cDNA; 407 BP.  
XX AC ABX43656;  
XX  
XX 21-FEB-2003 (first entry)  
XX DE Bovine EST associated with lactation/muscle/fat deposition #8821.  
XX  
XX Bovine, ss; EST; expressed sequence tag; lactation; LMFD;  
KW muscle deposition; fat deposition; genome mapping; gene identification;  
KW gene analysis; cattle breeding.  
XX  
OS Bos Taurus.

XX US2002137139-A1.

26-SEP-2002.

24-SEP-2001; 2001US-0960352.

12-JAN-1999; 99US-115707P.

11-JAN-2000; 2000US-0480902.

(BYAT// BYATT J C.

(MATH// MATHIALAGAN N.

(TAON// TAO N.

(WARR// WARREN W C.

Byatt JC, Mathialagan N, Tao N, Warren WC;

WPI; 2003-110599/10.

New nucleic acid associated with lactation, and muscle and fat  
deposition, useful for genome mapping, gene identification and  
analysis, cattle breeding, or for genetically improving cattle

Claim 2; SEQ ID No 8821; 245pp; English.

The invention relates to a purified nucleic acid molecule associated with  
lactation or muscle and fat deposition (designated LMFD), derived  
from cattle, and the LMFD nucleic acid can specifically hybridise to a  
second nucleic acid molecule comprising any of 15112 nucleotide  
sequences, appearing as ABX34836-ABX49947, or complements of them.  
Also included are: (1) a transformed cell having a nucleic acid  
comprising an LMFD nucleic acid linked to a promoter and a 3' non-  
translated sequence that functions in the cell to cause termination of  
transcription and addition of polyadenylated ribonucleotides to a 3' end  
of the mRNA molecule; and (2) determining a level or pattern of a  
nucleic acid in a bovine cell or tissue comprising: (a) incubating a marker  
nucleic acid (comprising any of the 15112 nucleic acid sequences or its  
complement or fragment) with a complementary nucleic acid molecule  
obtained from the bovine cell or tissue, where hybridisation between the  
marker nucleic acid and the complementary nucleic acid permits the  
detection of the molecule; and (b) detecting the level or pattern of the  
complementary nucleic acid, where the detection of the complementary  
nucleic acid is predictive of the level or pattern of the molecule.  
The LMFD nucleic acid is used for determining a level or pattern  
of a molecule in a bovine cell or tissue. It is useful for genome  
mapping, gene identification and analysis, cattle breeding, preparation  
of constructs for use in cattle gene expression, or for genetically  
improving cattle. The present sequence is one of the 15112 bovine  
LMFD EST (expressed sequence tag) nucleic acids.  
Note: The present sequence was not shown in the specification but  
was obtained in electronic format from the USPTO web site:  
seqdata.uspto.gov/sequence.html?DocID=20020137139.

Sequence 407 BP; 108 A; 101 C; 89 G; 109 T; 0 other;

Query Match 11.2%; Score 337.6; DB 25; Length 407;  
Best Local Similarity 90.9%; Pred. No. 5.2e-94;  
Matches 371; Conservative 0; Mismatches 34; Indels 3; Gaps 1;  
QY 1670 GGTAGACACAGTCCATTAAAGTCCAGAAATGCAACAGATGAATCCTAGAGGAGG 1729  
DB 3 GATACAGACAGTCCATTAAAGTCCAGAAATGCAACAGATGAATCCTAGAGGAGG 59  
QY 1730 TTGATTGCAAAATCTGGCTGAGCATATCTATTCACTGCTAGCAAGTCTGTGCCATTATG 1789  
DB 60 CTGATTGCAATCTGGCTGAGCACAATCTCTACTGCTAGCAAGTCTGTGCTATTATG 119  
QY 1790 TCCACACACATTTGGTGGTGGCTGCTCTCTACAGACACAGGAGGAAATGATCTCTCC 1849  
DB 120 TCCAGGCACATCTGGCTGAGCACAATCTCTCTCTCTACAGACACAGGAGGAAATGCTCTCC 179  
QY 1850 ACATTGTCGAGAACTCTTTTGTGATGAAAGAGGAGTCTGCTGCTGATTTTGCACCTG 1909

Db 180 ACATTGGTGAAGACTTCTTTGTGATGAAGAGGAAAGTCTCGTCTGTGATTTTGACTTG 239  
QY 1910 GGGTTCTCAGAAATTCAGAAAGATGTAGTAATCATGCCATACAGCTGCTGGGAAATTGT 1969  
Db 240 GGTTCCTCAGAAATTCAGAAAGATGTAGTATCATGCCATACAGCTTCTGGGAAATTGT 299  
QY 1970 GTCAATATCCACACTAGCAGGAACGATGAGTTTTTTATCACCCCGCAGCACAACCTGTC 2029  
Db 300 ATCACAATCACCCACAGCAGCAAGATGATGAGTTTTTTTATTACTCTAGCACAACCTATC 359  
QY 2030 CCATCAGTCTTCGAACTCAACTCTTACAGCAATGGGTACTTTCATGTC 2077  
Db 360 CCATCAGTCTTTGAACTCAACTCTTACAGCAATGGGTGCTCCATGTC 407

RESULT 12

ABX35577/c  
ID ABX35577 standard; cDNA; 418 BP.  
XX AC ABX35577;  
XX AC  
XX 20-FEB-2003 (first entry)  
XX DE Bovine EST associated with lactation/muscle/fat deposition #742.  
XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
KW muscle deposition; fat deposition; genome mapping; gene identification;  
KW gene analysis; cattle breeding.  
XX OS Bos Taurus.  
XX PN US2002137139-A1.  
XX PD 26-SEP-2002.  
XX PF 24-SEP-2001; 2001US-0960352.  
XX PR 12-JAN-1999; 99US-115707P.  
XX PR 11-JAN-2000; 2000US-0480902.  
XX PA (BYAT/) BYATT J C.  
PA (MATH/) MATHIALAGAN N.  
PA (TAON/) TAO N.  
PA (WARR/) WARREN W C.  
XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;  
XX WPI; 2003-110599/10.

XX New nucleic acid associated with lactation, and muscle and fat  
XX deposition, useful for genome mapping, gene identification and  
XX analysis, cattle breeding, or for genetically improving cattle  
XX  
XX Claim 2; SEQ ID No 742; 245pb; English.

XX The invention relates to a purified nucleic acid molecule associated with  
XX lactation or muscle and fat deposition (designated LMFD), derived  
XX from cattle, and the LMFD nucleic acid can specifically hybridise to a  
XX second nucleic acid molecule comprising any of 1512 nucleotide  
XX sequences, appearing as ABX34836-ABX49947, or complements of them.  
XX Also included are: (1) a transformed cell having a nucleic acid  
XX comprising an LMFD nucleic acid linked to a promoter and a 3' non-  
XX translated sequence that functions in the cell to cause termination of  
XX transcription and addition of polyadenylated ribonucleotides to a 3' end  
XX of the mRNA molecule; and (2) determining a level or pattern of a  
XX molecule in a bovine cell or tissue comprising: (a) incubating a marker  
XX nucleic acid (comprising any of the 1512 nucleic acid sequences or its  
XX complement or fragment) with a complementary nucleic acid molecule  
XX obtained from the bovine cell or tissue, where hybridisation between the  
XX marker nucleic acid and the complementary nucleic acid permits the  
XX detection of the molecule; and (b) detecting the level or pattern of the  
XX complementary nucleic acid, where the detection of the complementary  
XX nucleic acid is predictive of the level or pattern of the molecule.

CC The LMFD nucleic acid is used for determining a level or pattern  
CC of a molecule in a bovine cell or tissue. It is useful for genome  
CC mapping, gene identification and analysis, cattle breeding, preparation  
CC of constructs for use in cattle gene expression, or for genetically  
CC improving cattle. The present sequence is one of the 1512 bovine  
CC LMFD EST (expressed sequence tag) nucleic acids.  
CC Note: The present sequence was not shown in the specification but  
CC was obtained in electronic format from the USPTO web site:  
CC seqdata.uspto.gov/sequence.html?DocID=20020137139.

XX SQ Sequence 418 BP; 97 A; 116 C; 102 G; 102 T; 1 other;

Query Match 10.1%; Score 302; DB 25; Length 418;  
Best Local Similarity 84.1%; Pred. No. 6.4e-83;  
Matches 354; Conservative 0; Mismatches 61; Indels 6; Gaps 1;

QY 2112 ATGCAGTTCTGAACAGAGGGGACTGGGGGTCCCACTAGCACCCCACTACCTGATCA 2171  
Db 418 ATGCAGTTCTGAAGAGAGGGGCGCNCAGGAGGACC-----CGCGTCTCCCACTGGTCA 365  
QY 2172 GCCAGGAGCAGCTGGTGGGAAGCGCGCAGCCTGTGCTTCTTCCCAATGAAGCA 2231  
Db 364 GCCAGGAGCAGCTGGTGGCAAGGCTGCCAGGCTGTCTATCTCTCCAATGAAGCA 305  
QY 2232 CCATCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAACAGTAGGAAAGTTTA 2291  
Db 304 CCATCTCTCTCCCTGCCAGACCTTTTACCAAAATTTGCCATGAACAGTGGCGGTTTA 245  
QY 2292 TCCAGTATGGCATTTCTACAGTGGCAGACGATGACCAGGAATATCAGTCTCTAGTC 2351  
Db 244 TCCAGTATTTGCAATTTTATAGTTGCCAGCAAGATAATCAGGAAGATATCAGTCTTGT 185  
QY 2352 TTGCTGAGCAGCTGGGACAGAAAGCTTCCAGAACCTTTGTCTTGGGAAGTGAAG 2411  
Db 184 CTTCGAGCAGCAGTTGGACAAGAGGTTCCGGAACCTTTGTCTTGGGAAGTGAAG 125  
QY 2412 AAGATCAACACAGTGAATTTGGGAGGAAACAGCAGATTTGCTACCTGAAGGTGAGCAAT 2471  
Db 124 AAGATGAACACAGCGATTTTGTGAGGAGCAACGAGATTGCTACCTGAAGGTGAGCAAT 65  
QY 2472 CCAAGGAGCAGCAGATTTTATCAGTCTTTCACAGACTCCTTGGGCTTTGCTGGAGG 2531  
Db 64 CCAAGGAGCAGCAGATTTTATCAGTCTTTCACAGAGGCTCCTCGGGCTCTGCTGGAGG 5  
QY 2532 C 2532  
Db 4 C 4

RESULT 13

AAC05673  
ID AAC05673 standard; cDNA; 271 BP.

XX AC AAC05673;  
XX AC  
XX 06-OCT-2000 (first entry)  
XX DT  
XX DE Human secreted protein 5' EST, SEQ ID NO: 9748.  
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX OS Homo sapiens.  
XX PN EP1033401-A2.  
XX PD 06-SEP-2000.  
XX PF 21-FEB-2000; 2000EP-0200610.  
XX PR 26-FEB-1999; 99US-0122487.  
XX PA (GEST ) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
XX  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
XX Claim 1; SEQ ID 9748; 71pp + CD-ROM; English.  
XX  
XX The present sequence is one of a large number of 5' ESTs derived from  
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively  
XX identified within the present sequence. The 5' ESTs were prepared from  
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
XX sequences usually correspond mainly to the 3' untranslated region (UTR)  
XX of the mRNA because they are often obtained from oligo-dT primed cDNA  
XX libraries. Such ESTs are not well suited for isolating cDNA sequences  
XX derived from the 5' ends of mRNAs and even in those cases where longer  
XX cDNA sequences have been obtained, the full 5' UTR is rarely included.  
XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX They are used to obtain upstream regulatory sequences and to design  
XX expression and secretion vectors.  
XX  
XX Sequence 271 BP; 69 A; 77 C; 64 G; 61 T; 0 other;  
XX  
XX Query Match 8.2%; Score 246.4; DB 21; Length 271;  
XX Best Local Similarity 99.6%; Pred. No. 9.7e-66;  
XX Matches 247; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX 12 GCTGTCGGGGGGAAGTCTTTCTGAGGTTACTGTGGAGCACCACCAAGTCTGTGACCTCT 71  
XX  
XX 24 GCTGTCGGGGGGAAGTCTTTCTGAGGTTACTGTGGAGCACCACCAAGTCTGTGACCTCT 83  
XX  
XX 72 GGCCCTGCAACAGCAGCACCAGAGAAACAGACCTTGTATTATCCACAGCCTGGGAC 131  
XX  
XX 84 GGCCCTGCAACAGCAGCACCAGAGAAACAGACCTTGTATTATCCACAGCCTGGGAC 143  
XX  
XX 132 TGCTCTCTCCAGAGTCTCATCAGCTTTCCTAATCGACTGATGGAAATATTCCTCAAA 191  
XX  
XX 144 TGCTCTCTCCAGAGTCTCATCAGCTTTCCTAATCGACTGATGGAAATATTCCTCAAA 203  
XX  
XX 192 CACCACCAAGTCAAGGATACAGGACAGCGGCTCCCTGTTGTATGACATTTCTGCACC 251  
XX  
XX 204 CACCACCAAGTCAAGGATACAGGACAGCGGCTCCCTGTTGTATGACATTTCTGCACC 263  
XX  
XX 252 CGAAACTG 259  
XX  
XX 264 CGAAACCG 271  
XX  
XX  
XX RESULT 14  
XX ABX42469  
XX ID ABX42469 standard; cDNA; 387 BP.  
XX  
XX AC ABX42469;  
XX  
XX XX  
XX DT 20-FEB-2003 (first entry)  
XX  
XX DE Bovine EST associated with lactation/muscle/fat deposition #7634.  
XX  
XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
XX muscle deposition; fat deposition; genome mapping; gene identification;  
XX KW gene analysis; cattle breeding.  
XX  
XX OS Bos Taurus.  
XX  
XX OS  
XX PN US2002137139-A1.  
XX  
XX PD 26-SEP-2002.  
XX  
XX

PF 24-SEP-2001; 2001US-0960352.  
XX  
XX PR 12-JAN-1999; 99US-115707P.  
XX PR 11-JAN-2000; 2000US-0480902.  
XX  
XX PA (BYAT/) BYATT J C.  
XX PA (MATH/) MATHIALAGAN N.  
XX PA (TAON/) TAO N.  
XX PA (WARR/) WARREN W C.  
XX  
XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;  
XX WPI; 2003-110599/10.  
XX  
XX New nucleic acid associated with lactation, and muscle and fat  
XX deposition, useful for genome mapping, gene identification and  
XX analysis, cattle breeding, or for genetically improving cattle -  
XX  
XX Claim 2; SEQ ID No 7634; 245pp; English.  
XX  
XX The invention relates to a purified nucleic acid molecule associated with  
XX lactation or muscle and fat deposition (designated LMFD), derived  
XX from cattle, and the LMFD nucleic acid can specifically hybridise to a  
XX second nucleic acid molecule comprising any of 15112 nucleotide  
XX sequences, appearing as ABX34836-ABX49947, or complements of them.  
XX Also included are; (1) a transformed cell having a nucleic acid  
XX comprising an LMFD nucleic acid linked to a promoter and a 3' non-  
XX translated sequence that functions in the cell to cause termination of  
XX transcription and addition of polyadenylated ribonucleotides to a 3' end  
XX of the mRNA molecule; and (2) determining a level or pattern of a  
XX molecule in a bovine cell or tissue comprising: (a) incubating a marker  
XX nucleic acid (comprising any of the 15112 nucleic acid sequences or its  
XX complement or fragment) with a complementary nucleic acid molecule  
XX obtained from the bovine cell or tissue, where hybridisation between the  
XX marker nucleic acid and the complementary nucleic acid permits the  
XX detection of the molecule; and (b) detecting the level or pattern of the  
XX complementary nucleic acid, where the detection of the complementary  
XX nucleic acid is predictive of the level or pattern of the molecule.  
XX The LMFD nucleic acid is used for determining a level or pattern  
XX of a molecule in a bovine cell or tissue. It is useful for genome  
XX mapping, gene identification and analysis, cattle breeding, preparation  
XX of constructs for use in cattle gene expression, or for genetically  
XX improving cattle. The present sequence is one of the 15112 bovine  
XX LMFD EST (expressed sequence tag) nucleic acids.  
XX Note: The present sequence was not shown in the specification but  
XX was obtained in electronic format from the USPTO web site:  
XX seqdata.uspto.gov/sequence.html?DocID=20020137139.  
XX  
XX SQ Sequence 387 BP; 87 A; 97 C; 113 G; 90 T; 0 other;  
XX  
XX Query Match 8.2%; Score 244.8; DB 25; Length 387;  
XX Best Local Similarity 79.6%; Pred. No. 3.9e-65;  
XX Matches 301; Conservative 0; Mismatches 74; Indels 3; Gaps 1;  
XX  
XX QY 2614 CCTAATAACAGAGAACAGAGAAATGTTGCAGTATATGCTGAGAGTGCCATATTGTCT 2673  
XX  
XX Db 1 CCTGATCACGAGAACCGAGAGAGGGTTCGGTGTACGCTGAGAGCGCCATTACTGTCT 60  
XX  
XX QY 2674 TGTGAAGATGCTGTGAAATGTTAAGGATATTGGGGTTTTCAAGGAGACCAACAAA 2733  
XX  
XX Db 61 CGTGAAGATGCTGTGAAAGACGTTAAGGATATCGGGGTTTTCAAGGAGACCAACAAA 120  
XX  
XX QY 2734 GAGAGTCTGTTTTAGAACCTGACGACACCTTTTCTACCTCAATGCAACCGCAAAACT 2793  
XX  
XX Db 121 GAGAGTGTGTTTTAGAACCTGACGACACCTTTTCTACCTCAATGCAACCGCAAAACT 180  
XX  
XX QY 2794 TCTAGAATATATTCTGAGTGTGTTGGTGTAGTAAACGTTGGCAATG 2853  
XX  
XX Db 181 CCTGATATATATTCTGAGTGTGTTGGTGTAGTAACTTCGGGCACTCTGCAAGTG 240  
XX  
XX QY 2854 AAGTTCATGAGATGAGTTCCTTGTAGGTACAGCTTCTGGCTCAAGAGTTTGAAGGTGCC 2913  
XX  
XX Db 241 AAGGAGCGGCGCGGTTCTCTCGTAGG--CCCGCGCTGTGGCACGAGAGTGGACGGCGC 297  
XX



Result No.	Score	Query		DB	ID	Description	
		Match	Length				
1	53	1.8	2598	4	US-09-252-991A-4323	Sequence 4323, Appl	
C	53	1.8	2874	4	US-09-252-991A-4533	Sequence 4533, Appl	
3	46.2	1.5	2589	4	US-09-328-352-2160	Sequence 2160, Appl	
C	44.8	1.5	1830121	4	US-09-557-884-1	Sequence 1, Appl	
C	5	44.8	1830121	4	US-09-643-990A-1	Sequence 1, Appl	
C	6	38.6	1.3	1664976	4	US-08-916-421B-1	Sequence 1, Appl
7	37.2	1.2	1578	4	US-09-328-352-2557	Sequence 2557, Appl	
8	36.6	1.2	2790	3	US-08-895-601-2	Sequence 2, Appl	
C	35.6	1.2	7218	1	US-08-322-463-14	Sequence 14, Appl	
C	10	35.2	1.2	1032	4	US-09-328-352-2440	Sequence 2440, Appl
C	11	34.8	1.2	202001	4	US-09-734-674-3	Sequence 3, Appl
C	12	34.2	1.1	825	4	US-09-328-352-3343	Sequence 3343, Appl
13	34	1.1	1711	1	US-08-749-903-2	Sequence 2, Appl	
14	34	1.1	1711	4	US-09-088-641-2	Sequence 2, Appl	
C	15	34	1.1	9046	1	US-08-327-536-1	Sequence 1, Appl
C	16	34	1.1	9046	5	PCT-US95-04682-1	Sequence 1, Appl
17	33.6	1.1	2774	3	US-08-466-548B-1	Sequence 1, Appl	
18	33.6	1.1	2774	3	US-07-998-228F-1	Sequence 1, Appl	
19	33.6	1.1	2774	3	PCT-US93-12560-1	Sequence 1, Appl	
20	33.6	1.1	4821	3	US-08-913-374-1	Sequence 1, Appl	
C	21	33.4	1.1	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C	22	33.4	1.1	4401529	3	US-09-103-840A-1	Sequence 1, Appl
C	23	33.2	1.1	1534	4	US-08-858-207A-93	Sequence 93, Appl
C	24	33.2	1.1	1848	4	US-09-134-001C-447	Sequence 447, Appl
25	33.2	1.1	7672	4	US-09-220-132-24	Sequence 24, Appl	
26	33	1.1	3752	4	US-08-961-527-208	Sequence 208, Appl	
27	33	1.1	319608	4	US-09-539-333D-1	Sequence 1, Appl	

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4533
; LENGTH: 2874
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-4533

Query Match      1.8%; Score 53; DB 4; Length 2874;
Best Local Similarity 59.7%; Pred. No. 3.5e-06;
Matches 89; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 1036 CCATATTGACTATCTGCTCACTTTCATCTTCTGCGCAATCAACAGCACCATA 1095
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1918 CCATCGGACTACTGCTGCTCTCCCTACCTGCTGTTCCGCAACGGCTGACCCGCGCA 1859
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1096 CATTGCTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTTGATCCATAAGCTTGG 1155
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1858 CATCGCGCGGATCACTGACATGCGGTGATCGGCTCGATCTCTGCGCGCGCGG 1799
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1156 GGCCTTCTTCATACGAGGCTCGATG 1184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1798 GGCCTTCTTCATGCGCGGAGCTTCAAG 1770
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-09-328-352-2160
; Sequence 2160, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2160
; LENGTH: 2589
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-2160

Query Match      1.5%; Score 46.2; DB 4; Length 2589;
Best Local Similarity 55.2%; Pred. No. 0.00053;
Matches 90; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 1011 TTCTGTTTCTACCACTTATAGATCCATATTTGACTATCTGCTGCTCACTTTCTCTCT 1070
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 965 TTGTTTACACACCATGTCACCGTAGGCACATGACTACTTACTGCTTATGTGATTT 1024
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1071 TCTGCCATAACATCAAGCACCACATACATCTGTTTCAGGCAATAATCTCAACATCCCAATCT 1130
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1025 ATAAAGTGGCTTGATGTTCCGTATATGCTGAGGAGACACCTAACTTACCTTTCG 1084
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1131 TCAGTACCTTGATCCATAAGCTTTGGGGCTTTCTTCATACGAC 1173
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1085 TTGGTCAGCTTTTACGTGGTGGCGGTGCTCTTTTATTCGCCG 1127
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
; US-09-557-884-1

Query Match      1.5%; Score 44.8; DB 4; Length 1830121;
Best Local Similarity 45.2%; Pred. No. 0.19;
Matches 227; Conservative 0; Mismatches 257; Indels 18; Gaps 1;

QY 1011 TTCTGTTTCTACCACTTATAGATCCATATTTGACTATCTGCTGCTCACTTTCTCTCT 1070
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 808330 TTGTTTATGTGCTTGGCCACCGTAGTCATATCGATTACTTATTTGCTTTCTTATGTGCTTT 808271
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1071 TCTGCCATAACATCAAGCACCACATACATGCTTTCAGGCAATAATCTCAACATCCCAATCT 1130
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 808270 ATCATCAAGTCTTGTTCGCCACATATCGCGCAGGATTAATTAATTTCTGGCCTA 808211
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1131 TCAGTACCTTGATCCATAAGCTTTGGGGCTTCTTCATACGACGAGGCTCGATGAACAC 1190
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 808210 TAGGTAGAATGTTTCGTAGTTGGGGGCACTTCTTTATTCGCCGCACTTTTAAAGGGAATC 808151
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1191 CAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATAGTTGAATTAC 1250
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 808150 GA-----CTATATCTGCCAATTTTCGCGAATATTTATCGAATCTAT 808109
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1251 TTCGACAGCAGCAATTTCTGGAGATCTTCTCGAAGGCACACGCTTAGAGTGGAAAAA 1310
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 808108 TCCATCGAGGCTATTCCTGTCGAATATTTTATGAGGCGCGTCTGTTCTGTCAGGTCGTT 808049
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1311 CCTCTTGTCTCGGCGAGGACTTTTGTCTAGTTGTGTAGATATCTCTGTCTTCAATGTCA 1370
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 808048 TGTGSCACCAAAAAACAGGTATGATGTCATGACACTTCAAGCATTTGCAACATAGTCAAA 807989
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1371 TCCGACATCTTGATTAATACCTGTTGGAATCTCTTATGATCGCATATTCGAAGTCACT 1430
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 807988 CTCGCCCAATTTTCGATTTGTTCTCTGTTTACGTTGGTTTGAACACACATTTAGAGTAGACA 807929
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1431 ACAATGGTGAACAACCTGGGCAAAACCTTAAGAAGATGAGAGCCTGTGGAGTGTAGCAAGAG 1490
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 807928 CTTATGCTAAGGAATTCGTGGTGCAGCGAAAGAAAAAATGCGGGTTTACTTCTC 807869
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1491 GTGTTATTAGAATGTTTACGAAA 1512
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



Db 807868 GTGTAATTAAAAAATTACGTAA 807847

RESULT 5

US-09-643-990A-1/c

; Sequence 1, Application US/09643990A

; Patent No. 6528289

GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann.

; Mark D. Adams

; Owen White

; Hamilton O. Smith

; J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of

the Haemophilus influenzae Rd Genome, Fragments

Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville,

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/643,990A

FILING DATE: 23-Aug-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,429

FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787

FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB186P1C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-610-5790

TELEFAX: 310-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-643-990A-1

Query Match 1.5%; Score 44.8; DB 4; Length 1830121;  
Best Local Similarity 45.2%; Pred. No. 0.19;  
Matches 227; Conservative 0; Mismatches 257; Indels 18; Gaps 1;

Qy 1011 TTCGTTTCTACCACTTCATAGATCCCATATTGATCTCTGCTGCTCACTTTCATCTCT 1070

Db 808330 TTGTTATTGTGCTTTCGCCACCGTAGCTATCTGATTTATGCTTTCTTTATGTGCTTT 808271

Qy 1071 TCTGCCATAACCAAGCACCATATGCTTCAGGCAATAATCTCAACATCCCAATCT 1130

Db 808270 ATCATCAAGGCTTTGTCGCCACATATCGCGGCGAGGATTAATCTTAATTTCTGCGCTA 808211

Qy 1131 TCAGTACCTTCATCCATAAGCTTTGGGGCTTCTTCATACGACGAAGGCTCGATGAACAC 1190

Db 808210 TAGGTAGATGTTTCGTAGTTGGGGGCGATCTTTATTTCGCCGACATTTTAAGGAATC 808151

Qy 1191 CAGATGGACGGAAGATGTTCTCTATAGAGCTTTCTCATAGAGCTTTCTCATGGGCAATAGTGAATPAC 1250

Db 808150 GA-----CTATATTCTGCATTTTTCGCCAATATTTATCAGAATCTAT 808109

Qy 1251 TTCACACAGCAAAATCTTTGGAGATCTTCTCGAAGGCACACAGTCTTAGGAGTGGAAAA 1310  
Db 808108 TCCATCGAGGCTATTTCGTCGAATATTTATTAGGGGGGCTGTTCTCGTACAGTCCGT 808049  
Qy 1311 CCTCTTGTGTCGGGCGAGGACTTTTTCAGTTGTTGGTAGATACTCTGTCTACCAATGTCA 1370  
Db 808048 TGCTGGCACCAAAAAACAGGTATGATGTCAATGACACTTCAAGCATTTGCAACATAGTCAA 807989  
Qy 1371 TCCCAGACATCTTGATAATACCTGTGGAATCTCTATGATCGCATTTACGAAGTCACT 1430  
Db 807988 CTGCCCCCAATTCGATTTGTTCTCTGTTTACGTTGGTTTGAACACGTATTAGAAGTAGACA 807929  
Qy 1431 ACAATGGTGAACAACTGGGCAAACTTAAGAAAGATGAGAGCTGTGGAGTGTAGCAAGAG 1490  
Db 807928 CTTAATGCTAAGGAATTACGTGTGTCAGCGAAGAAAAAGAAATCGCGGTTTAGTACTTC 807869  
Qy 1491 GTGTTATTAGAAATGTTAGAAA 1512  
Db 807868 GTGTAATTAAAAAATTACGTAA 807847

RESULT 6

US-08-916-421B-1/c

; Sequence 1, Application US/08916421B

; Patent No. 6503729

GENERAL INFORMATION:

APPLICANT: Bult et al.

TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus

Patent No. 6503729

TITLE OF INVENTION: jannaschii

FILE REFERENCE: PB275

CURRENT APPLICATION NUMBER: US/08/916,421B

CURRENT FILING DATE: 1997-08-22

PRIOR APPLICATION NUMBER: US 60/024,428

PRIOR FILING DATE: 1996-08-22

NUMBER OF SEQ ID NOS: 3

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 1664976

TYPE: DNA

ORGANISM: Methanococcus jannaschii

FEATURE:

NAME/KEY: misc feature

LOCATION: (28222)..(28222)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (28257)..(28258)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (84773)..(84773)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (84808)..(84808)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (84812)..(84812)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (98120)..(98120)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (98159)..(98159)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (98239)..(98239)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (98266)..(98266)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature

LOCATION: (98343)..(98343)

OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature  
LOCATION: (103998)..(103998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (148948)..(148948)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (163385)..(163385)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191989)..(191989)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191995)..(191995)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (231980)..(231980)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234220)..(234220)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature

LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1470091)..(1470091)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1664855)..(1664855)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-421B-1

Query Match 1.3%; Score 38.6; DB 4; Length 1664976;

Best Local Similarity 50.8%; Pred. No. 18;

Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 678 TTTTATTCAAGAGCGAGATGTGCATAAGGGCATGTTGCCACCAATGTGACTGAAATG 737  
|||||  
Db 541143 TCTTTATTAAACATAGAGAGGATAACAGATGCATGATATAATAGTAGTGAATAGTGCATG 541084  
  
QY 738 TCGTGAACAGCAGTAGAGTACAGAGGCAATGTCAGAGTGGCTGCTGAATTAACCCCTG 797  
|||||  
Db 541083 CTCTTAAATGAAATATTAAAGAAATAACAAAATTTGTCAGTTGTAGGAGAGCCAA 541024  
  
QY 798 ATGGTTCTGCCAGCAGCAATCAAGCCGTTTAAACAAAGCTAAAGAAAGAAAGCTAAAGGA 857  
|||||  
Db 541023 AAACCTCTGAAAGTTGGATACATAATGTTGGTTAATGTAAGAGCATGTGAACACTA 540964  
  
QY 858 T 858

Db 540963 T 540963

RESULT 7

US-09-328-352-2557  
; Sequence 2557, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 2557  
; LENGTH: 1578  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-2557

Query Match 1.2%; Score 37.2; DB 4; Length 1578;  
Best Local Similarity 54.3%; Pred. No. 0.32;  
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;  
  
Qy 728 ACTGAAATGTGCTGAACAGCAGTAGTACAGAGGCAATTCAGAGTGGCTGCTGAA 787  
Db 1189 ATTGAAATGTACCGAAGCAGTCTTCTATGCTAGGCGCAATGTCTTGTTTATGACTCA 1248  
  
Qy 788 TTAACCTGTGTTCTGCCAGCAGCAATCAAAAGCGTTTAAACAAGTGAAGAAA 847  
Db 1249 TTCATGTGTTCTGCTGAAGCCCACTGCAATATATTGTAGATGTTTATTGCAATGAAA 1308  
  
Qy 848 GCTAAAGGATTCTTCAA 865  
Db 1309 GCGAAGGTATTGCTAAA 1326

RESULT 8

US-08-895-601-2  
; Sequence 2, Application US/08895601  
; Patent No. 6060262  
; GENERAL INFORMATION:  
; APPLICANT: Beer-Romero, Peggy  
; APPLICANT: Strack, Peter J.  
; APPLICANT: Glass, Susan J.  
; APPLICANT: Rolfe, Mark  
; TITLE OF INVENTION: REGULATION OF KAPPA B (IKB) DEGRADATION,  
; TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/895,601  
; FILING DATE: 16-JUL-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MIV-096.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000

; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2790 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2..2782  
US-08-895-601-2

Query Match 1.2%; Score 36.6; DB 3; Length 2790;  
Best Local Similarity 55.0%; Pred. No. 0.77;  
Matches 72; Conservative 0; Mismatches 59; Indels 0; Gaps 0;  
  
Qy 735 ATGTGCTGAACAGCAGTAGTACAGAGGCAATTCAGAGTGGCTGCTGAATTAACC 794  
Db 531 ATTTACCTAAACCAAGTGGCTCAGAAAGATGATAATGAGAACAGGCTGAGGAATTAGAGC 590  
  
Qy 795 CTGATGTTCTGCCAGCAGCAATCAAAAGCGTTTAAACAAGTGAAGAAAAGCTAAAA 854  
Db 591 CTGGCTGGTGTGTTTGGACCAACACAGATGCTGTGTCATTTCAGCAACAACAAGAAC 650  
  
Qy 855 GGATTCTTCAA 865  
Db 651 CTTCTCCTCTA 661

RESULT 9

US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single



```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/088,641
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/749,903
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0163 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1711 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6312895e
IMMEDIATE SOURCE:
CLONE: 989953
US-09-088-641-2

Query Match 1.1%; Score 34; DB 4; Length 1711;
Best Local Similarity 63.4%; Pred. No. 3.8;
Matches 52; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1598 GCTCTACTTTCCCTGGAGCAAGCGTTGTACACGAGCTATACCTCTTCAAGACCCAGTGAT 1657
|||||
DB 1270 GCTGTACAGTCCTCGGCAAGCAGGTTTACCTGTATCTCATCGGAAGGCTCTGTGAT 1329
|||||

QY 1658 GCTGCTGATGAAGGTAGACACA 1679
|||||
DB 1330 GCTGCAGGTTGATGTAGACACA 1351
|||||

RESULT 15
US-08-227-536-1/c
Sequence 1, Application US/08227536
Patent No. 5658784
GENERAL INFORMATION:
APPLICANT: Eckner, Richard
APPLICANT: Ewen, Mark
APPLICANT: Livingston, David
TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
TITLE OF INVENTION: FACTOR P300 AND USES OF P300
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,536
FILING DATE: 14-APR-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Williams Ph.D., Kathleen A.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: DPCT-308XX

```

**TELECOMMUNICATION INFORMATION:**

TELEPHONE: (617) 542-2290

TELEPHONE: (617) 342-2230  
TELEFAX: (617) 451-0313

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 9046 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

**FEATURE:**

NAME/KEY: CDS

LOCATION: 1200..8441

US-08-227-536-1

Query Match 1.1%; Score 34; DB 1; Length 9046;

Query Match	1.1%;	Score 34; DB
Best Local Similarity	61.1%;	Pred. No. 13;

```

DESC Local Similarity 01.1%; FREQ. NO. 15;
Matches 55; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

```

939 TCCTTTGGAAACATTCAAAATTCACAAAGGTCAACTTGAGATGGTTAAAGCTGCAACTGAGA 998

QY 359 1CTTTTGGACATCTCAATTTTCAATAGGTCACCTTGGATGGTTAAAGCTGCATATCTAGAG 358

Db 4235 TGTTCGGTAGATTCCATTTTACAGCTTCCACTTTAGACTCTGAATATATCTCCGGGTG 4176

Qy 999 CGAATTGCCGCTCTGTGTTCTACCAAGTTC 1028

[illegible]

Search completed: January 10, 2004, 06:15:21

Job time : 163.864 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: January 10, 2004, 04:06:03 ; Search time 695.063 Seconds  
(without alignments)  
15003.329 Million cell updates/sec

Title: US-09-935-290-1  
Perfect score: 3003  
Sequence: 1 ttccggaccaggctgtcg.....tctgagatcttccagagttt 3003

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2276164 seqs, 1736306516 residues  
Total number of hits satisfying chosen parameters: 4552328

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3001	99.9	3003	11	US-09-935-290-1
2	2970.2	98.9	4033	13	US-09-814-353-20793
3	2958	98.5	3210	13	US-10-288-252-26
4	2484	82.7	2484	11	US-09-935-290-3
5	1896.2	63.1	2646	10	US-09-917-800A-1648
6	581.6	19.4	612	10	US-09-833-381-1961
7	480	16.0	868	10	US-09-833-381-1958
8	437	14.6	482	11	US-09-918-995-21644
9	385.4	12.8	399	11	US-09-918-995-21644
10	348.4	11.6	440	10	US-09-833-381-1964
11	346.6	11.5	396	10	US-09-833-381-1959
12	342.2	11.4	429	10	US-09-960-352-7384
13	337.6	11.2	407	10	US-09-960-352-8821
14	302	10.1	418	10	US-09-960-352-742
15	268.6	8.9	379	13	US-09-814-353-16182

16	244.8	8.2	387	10	US-09-960-352-7634	Sequence 7634, Ap
17	233	7.8	303	10	US-09-960-352-13086	Sequence 13086, A
18	209.6	7.0	264	10	US-09-960-352-6239	Sequence 6239, Ap
19	207.2	6.9	244	13	US-09-814-353-3482	Sequence 3482, Ap
20	207.2	6.9	244	13	US-09-814-353-9798	Sequence 9798, Ap
21	171	5.7	501	10	US-09-960-352-4061	Sequence 4061, Ap
22	129	4.3	308	10	US-09-833-381-1963	Sequence 1963, Ap
23	109	3.6	2682	13	US-10-353-690-31	Sequence 31, Appl
24	109	3.6	2755	13	US-10-288-252-27	Sequence 27, Appl
25	93.6	3.1	399	10	US-09-960-352-54	Sequence 54, Appl
26	83.2	2.8	404	10	US-09-960-352-5756	Sequence 5756, Ap
27	75.2	2.5	467	13	US-10-027-632-297935	Sequence 297935, A
28	75.2	2.5	467	13	US-10-027-632-297936	Sequence 297936, A
29	75.2	2.5	467	14	US-10-027-632-297935	Sequence 297935, A
30	75.2	2.5	467	14	US-10-027-632-297936	Sequence 297936, A
31	75.2	2.5	634	13	US-10-027-632-41737	Sequence 41737, A
32	75.2	2.5	634	13	US-10-027-632-41738	Sequence 41738, A
33	75.2	2.5	634	14	US-10-027-632-41737	Sequence 41737, A
34	75.2	2.5	634	14	US-10-027-632-41738	Sequence 41738, A
35	62	2.1	442	10	US-09-833-381-1965	Sequence 1965, Ap
36	59.2	2.0	469	13	US-10-027-632-62086	Sequence 62086, A
37	59.2	2.0	469	13	US-10-027-632-62087	Sequence 62087, A
38	59.2	2.0	469	13	US-10-027-632-63434	Sequence 63434, A
39	59.2	2.0	469	13	US-10-027-632-63435	Sequence 63435, A
40	59.2	2.0	469	14	US-10-027-632-62086	Sequence 62086, A
41	59.2	2.0	469	14	US-10-027-632-62087	Sequence 62087, A
42	59.2	2.0	469	14	US-10-027-632-63434	Sequence 63434, A
43	59.2	2.0	469	14	US-10-027-632-63435	Sequence 63435, A
44	57.6	1.9	2262	12	US-10-369-493-33072	Sequence 33072, A
45	56.4	1.9	344	10	US-09-960-352-4542	Sequence 4542, Ap

ALIGNMENTS

RESULT 1  
US-09-935-290-1  
; Sequence 1, Application US/09935290  
; Publication No. US20030044948A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller-Libermann, Rosana  
; TITLE OF INVENTION: 56919, A NOVEL HUMAN ACYLTRANSFERASE AND USES THEREOF  
; FILE REFERENCE: NMI-186  
; CURRENT APPLICATION NUMBER: US/09/935,290  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/226,509  
; PRIOR FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3003  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (341)...(2827)  
; NAME/KEY: misc feature  
; LOCATION: (1)...(3003)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-935-290-1

Query Match	99.9%	Score	3001	DB	11	Length	3003
Best Local Similarity	100.0%	Pred. No.	0				
Matches	3003	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Oy	1	TTTCGGCACCAAGGCTGCTGCGGGGGGACTCTTTCTGAGGTACTCTGGAGCACCAAGTC	60				
Db	1	TTTCGGCACCAAGGCTGCTGCGGGGGGACTCTTTCTGAGGTACTCTGGAGCACCAAGTC	60				
Oy	61	TGTCAGGCTTGGCCGTGTCGAAACAGGACCCAGAGAACCAAGCTTGTATTACCCCA	120				
Db	61	TGTCAGGCTTGGCCGTGTCGAAACAGGACCCAGAGAACCAAGCTTGTATTACCCCA	120				



Qy	121	CAGCTGGGACTGTCTTCCAGAGTCTCCATCAGCTTTGCTAATCGACTGATTGGAAT	180
Db	121		
Qy	181	AATTCCTCAACACCAAGTCAAGATACAGGACAGCGGCTCCCTGTTGTATGGA	240
Db	181		
Qy	241	CATTCTGACCCGAACTGATAGCTGAGTCTGAGTCTTATGTTATGAACACAGAGAAC	300
Db	241		
Qy	301	TTTCATCCGACGACATGATTGGGAATTACACTTTGTGACATGGATGAATCTCGACTGAC	360
Db	301		
Qy	361	CTTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCAGATACAGTGTGGTGGATG	420
Db	361		
Qy	421	TAAGCACACAGTGAGGAAGGGTGAGTGGCTTTAGACCCACCGTCTTTCAGATCTGC	480
Db	421		
Qy	481	AACTTTAAATGGAAGAAAGCCTAATGAGTTCGAAAGGCCATTTGTTGGAAGATGTTG	540
Db	481		
Qy	541	TTACTCTGCACTCCCGAGAGTGGGACAAATTTTCAACCCCGATATCCCGTCTTTGGG	600
Db	541		
Qy	601	TTTGGCGAATGTTATTTATATCAATGAATCTACACAGACACCGCGGATGGCTGCAAG	660
Db	601		
Qy	661	ACGCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTCATAAGGGGCATGTTTGGCAC	720
Db	661		
Qy	721	CAATGTGACTGAAATGTGTGAACAGCAGTAGAGTACAAGAGCAATTCAGAAAGTGC	780
Db	721		
Qy	781	TGCTGAATTAACCTGATGGTCTGCCAGCAGCAATCAAGCCGTTAACAAGTGAA	840
Db	781		
Qy	841	AAAGAAAGCTAAAGGATCTTCAAGAAATGGTTGCCACTGTCTCACCGGCAATGATCAG	900
Db	841		
Qy	901	ACTGACTGGGTGGTGTCTTAAACTGTTCAACAGCTTCTTTTGGAAACATTCAAATTC	960
Db	901		
Qy	961	CAAGGTCACATTGAGATGTTAAAGCTGCAACTGAGACGAATTTGCCGCTCTGTTTCT	1020
Db	961		
Qy	1021	ACAGTTTCATAGATCCCATATTTGACTATCTGCTGCTCACATTTTCATTTCTCTGCCATA	1080
Db	1021		
Qy	1081	CATCAAGACCAATACATTGCTTCAGGCAATAATCTCAACATCCCAATCTTCAGTACCTT	1140
Db	1081		
Qy	1141	GATCCATAAGCTTTGGGGCTTCTTCATACAGAGGCTCGATGAACACACAGATGGAG	1200
Db	1141		
Qy	1201	GAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATAGTTGAATTTACTTCGACAGCA	1260
Db	1201		

Db	1201	GAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATAGTTGAATTTACTTCGACAGCA	1260
Qy	1261	GCAATTTCTGGAGATCTTCTCGAAGGCACAGTTCCTAGAGTGGAAACCTCTTGTGC	1320
Db	1261		
Qy	1321	TCGGGACAGACTTTTGTGTCAGTTGTGTGATAGTACTCTGTCTACCAATGTCTATCCAGACAT	1380
Db	1321		
Qy	1381	CTTGATAATPACCTGTTGGAACTCTCTATGATCGCAATTCGAAAGTTCACCAATGGTGA	1440
Db	1381		
Qy	1441	ACAACTGGGCAAACTTAAGAAGATGAGAGCTGTGGAGTGTAGCAGAGGTGTTATTAG	1500
Db	1441		
Qy	1501	AAATGTTACGAAAACTATGTTGTGTCGAGTGGATTTTGACAGCCATTTTCTTAA	1560
Db	1501		
Qy	1561	GGAAATATTTAGAAAGCCAAAGTCAGAAACCGGTGTCTCTACTTTCCCTGGAGCAAGC	1620
Db	1561		
Qy	1621	GTTGTTACAGCTATATCTTCTTCAAGACCCAGTGTCTGTCTGATGAAGGTAGAGACAC	1680
Db	1621		
Qy	1681	GTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGAGGTGATTGCAAA	1740
Db	1681		
Qy	1741	TCGGCTGAGCATAATTTCTATTCAGTCTAGCAAGTCTCTGTCATTTATGTCACACACAT	1800
Db	1741		
Qy	1801	TGTGGCTGCTGCTCTCTACAGACAGGACAGGAATGTCTCTCACATTTGTCGA	1860
Db	1801		
Qy	1861	AGACTTCTTTGTGATGAAGAGAGAGTCTGCTGCTGTGATTTTGACTCTGGGTTCTCAGG	1920
Db	1861		
Qy	1921	AAATTCAGAAAGATGTAGTAATGCATGCCATACAGCTGTGGGAAATTTGTGCACATCAC	1980
Db	1921		
Qy	1981	CCACACTAGAGAAAGATGATGTTTTTATACCCCCAGACAACTGTGCCATCAGTCTT	2040
Db	1981		
Qy	2041	CGAACTCAACTTCTACAGCAATGGGGAATTTATCATGCTTTTATCATGGAGCCATCATAGC	2100
Db	2041		
Qy	2101	TTGACGCTTTTATGCAATGTTGAAAGAGGGGACTGGGGGGTCCCACTAGCACCCACC	2160
Db	2101		
Qy	2161	TAACTGTATCAGCAGGAGCAGCTGTGGGGAAGGGGGCAGCTGTGCTACCTTCTCTC	2220
Db	2161		
Qy	2221	CAATGAAGGCACCATCTCACTGCTTGCAGACATTTTACCAGTCTGCCATGAACAGT	2280
Db	2221		
Qy	2281	AGGAAAGTTTATCCAGTATGGCATTTCTTACAGTGGCAGACGATGACCAGGAATAT	2340
Db	2281		

Db 2281 AGGAAGTTTATCCAGTATGCAATCTTACAGTGCAGACGACGATGACCGAGAGATAT 2340  
Qy 2341 CAGTCTAGTCTTGTGAGCAGCAGTGGGACAAAGAGCTTCAGAACCTTTGTCTTTGAG 2400  
Db 2341 CAGTCTAGTCTTGTGAGCAGCAGTGGGACAAAGAGCTTCAGAACCTTTGTCTTTGAG 2400  
Qy 2401 AAGTATGAAGAGATGAAGACAGTACTTTGGGGAGGAACAGCGAGATTTGCTACCTGAA 2460  
Db 2401 AAGTATGAAGAGATGAAGACAGTACTTTGGGGAGGAACAGCGAGATTTGCTACCTGAA 2460  
Qy 2461 GGTGAGCCAAATCCAGGAGCAGCAGCAGTATATACCTTTTACAGACACTCTTTGGGCC 2520  
Db 2461 GGTGAGCCAAATCCAGGAGCAGCAGCAGTATATACCTTTTACAGACACTCTTTGGGCC 2520  
Qy 2521 TTTGCTGAGGCTTACAGCTCTGCTGCCATCTTTGTTTCAACAATTTAGTGTCTGTTCC 2580  
Db 2521 TTTGCTGAGGCTTACAGCTCTGCTGCCATCTTTGTTTCAACAATTTAGTGTCTGTTCC 2580  
Qy 2581 AGAACCTGAGTATCTGCAAAAGTTGCAACAATACCTAATACCAAGAACAGAAATGT 2640  
Db 2581 AGAACCTGAGTATCTGCAAAAGTTGCAACAATACCTAATACCAAGAACAGAAATGT 2640  
Qy 2641 TGCAGTATATGCTGAGAGTGCCACATATTTGCTGTGAAGATCTGTGAAAATGTTTAA 2700  
Db 2641 TGCAGTATATGCTGAGAGTGCCACATATTTGCTGTGAAGATCTGTGAAAATGTTTAA 2700  
Qy 2701 GGATATTGGGTTTTCAAGGAGACAAACAAAAGAGTGTCTGTTTGAACCTGAGCAG 2760  
Db 2701 GGATATTGGGTTTTCAAGGAGACAAACAAAAGAGTGTCTGTTTGAACCTGAGCAG 2760  
Qy 2761 CACTTTTCTACTCAATGCAACCGCAAAAATCTTAGAATATATCTGAGTTTGTGGT 2820  
Db 2761 CACTTTTCTACTCAATGCAACCGCAAAAATCTTAGAATATATCTGAGTTTGTGGT 2820  
Qy 2821 GCTGTAGTAACTGTGCTGCTGCTGCAATGAGGTGATGAGTGTCTGTTGAGG 2880  
Db 2821 GCTGTAGTAACTGTGCTGCTGCTGCAATGAGGTGATGAGTGTCTGTTGAGG 2880  
Qy 2881 TACCAGCTTCTGGCTCAAGAGTTTGAAGTGTGCTTGCAGGGGTGAGGCTGCTGCTGNC 2940  
Db 2881 TACCAGCTTCTGGCTCAAGAGTTTGAAGTGTGCTTGCAGGGGTGAGGCTGCTGCTGNC 2940  
Qy 2941 CGAAGTATCTCTGGAAGACAAAGTCTTCTNCTCCATGATCTGAGATCTTCCCAGC 3000  
Db 2941 CGAAGTATCTCTGGAAGACAAAGTCTTCTNCTCCATGATCTGAGATCTTCCCAGC 3000  
Qy 3001 TTT 3003  
Db 3001 TTT 3003

RESULT 2  
US-09-814-353-20793  
; Sequence 20793, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Pamela  
; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
; FILE REFERENCE: MRI-006B  
; CURRENT APPLICATION NUMBER: US/09/814,353  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/191,031  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: US 60/207,124  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: US 60/211,940  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 60/216,820  
; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: US 60/220,661  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: US 60/257,672  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 22037  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20793  
; LENGTH: 4033  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1, 2, 4030, 4031, 4032, 4033  
; OTHER INFORMATION: n = A,T,C or G  
US-09-814-353-20793

Query Match 98.9%; Score 2970.2; DB 13; Length 4033;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2994; Conservative 0; Mismatches 5; Indels 2; Gaps 2;  
Qy 3 CGGCACCAGGCTGCTGGGGGGGACTCTTTCTGAGGTTTACTGTGAGCAGCACCACCAAGTCTG 62  
Db 3 CGGCACCAGGCTGCTGGGGGGGACTCTTTCTGAGGTTTACTGTGAGCAGCACCACCAAGTCTG 62  
Qy 63 TCAGCCTCTGGCCGTCAAAACAGGCACCCAGAGAACAGACCTTGTCTTATTCACCCACA 122  
Db 63 TCAGCCTCTGGCCGTCAAAACAGGCACCCAGAGAACAGACCTTGTCTTATTCACCCACA 122  
Qy 123 GCCTGGGACTGTCTTCTCCAGAGTCTCATCAGCTTTGCTTAATCGACTGATGGAATAA 182  
Db 123 GCCTGGGACTGTCTTCTCCAGAGTCTCATCAGCTTTGCTTAATCGACTGATGGAATAA 182  
Qy 183 TTCTCTAAACACCAACCAAGTCAAGGATACAGGACGAGCGCTCCCTGTTGTATGGACA 242  
Db 183 TTCTCTAAACACCAACCAAGTCAAGGATACAGGACGAGCGCTCCCTGTTGTATGGACA 242  
Qy 243 TTCTGACCCGAAACTGATAGTCTGAGTCTGAACTTTTATGTATGAAACAGAGAACTT 302  
Db 243 TTCTGACCCGAAACTGATAGTCTGAGTCTGAACTTTTATGTATGAAACAGAGAACTT 302  
Qy 303 TCATCCAGACACATGATTTGGGAATTAACATTTGTGACATGGAATCTGCACTGACCC 362  
Db 303 TCATCCAGACACATGATTTGGGAATTAACATTTGTGACATGGAATCTGCACTGACCC 362  
Qy 363 TTGTGTAAATAGATGTTTCTTATCTGACCATTCATCAGATACAGTGTGTTGTCGATGA 422  
Db 363 TTGTGTAAATAGATGTTTCTTATCTGCCAATTCATCAGATACAGTGTGTTGTCGATGA 422  
Qy 423 AGCACACAAGTGAAGAAATGGGTGAGTGTGGCTTTAGACCCACCGCTCTTCAGATCTGCAA 482  
Db 423 AGCACACAAGTGAAGAAATGGGTGAGTGTGGCTTTAGACCCACCGCTCTTCAGATCTGCAA 482  
Qy 483 CTTTAAATGGAAGAAAGCCTAATGAGTGGAAAGGCCATTTGTTGGAGATGTTGTT 542  
Db 483 CTTTAAATGGAAGAAAGCCTAATGAGTGGAAAGGCCATTTGTTGGAGATGTTGTT 542  
Qy 543 ACTCTGCACTCCCGAGCTGGGACAAATTTTCAACCCGAGTATCCCTCTTTGGTT 602  
Db 543 ACTCTGCACTCCCGAGCTGGGACAAATTTTCAACCCGAGTATCCCTCTTTGGTT 602  
Qy 603 TGGGAATGTTTATTTATCAATGAAACTCACACAGACACCGGGATGGCTTCCAGAG 662  
Db 603 TGGGAATGTTTATTTATCAATGAAACTCACACAGACACCGGGATGGCTTCCAGAG 662  
Qy 663 GCCTTTCTTACGTTCTTTTATTCAAGAGCAGATGTGCATAGGGCATGTTTCCACCA 722  
Db 663 GCCTTTCTTACGTTCTTTTATTCAAGAGCAGATGTGCATAGGGCATGTTTCCACCA 722  
Qy 723 ATGTGACTGAAATGTCTGAAACAGCAGTATGATACAGAGGCAATTCAGAGTGGCTG 782  
Db 723 ATGTGACTGAAATGTCTGAAACAGCAGTATGATACAGAGGCAATTCAGAGTGGCTG 782  
Qy 783 CTGAATTTAAACCCCTGATGTTCTGCCCCAGCAGCAATCAAAAGCCGTTTAAACAAAGTGA 842

Db 783 CTGAATTTAAACCCCTGATGGTTCTGCCAGCAGCAATCAAAAGCCGTTTAAACAAAGTGAAAA 842  
Qy 843 AGAAAGCTAAAGGATTTCTTCAAGAAATGTTGTCACCTGCTCACCAGCAATGATCAGAC 902  
Db 843 AGAAAGCTAAAGGATTTCTTCAAGAAATGTTGTCACCTGCTCACCAGCAATGATCAGAC 902  
Qy 903 TGAATGGGTGGGTGCTGCTTAAACCTGTTCAACAGCTTCTTTTGGAAATTTGCGGCTTCTGTTCTAC 962  
Db 903 TGAATGGGTGGGTGCTGCTTAAACCTGTTCAACAGCTTCTTTTGGAAATTTCAAAATTCACA 962  
Qy 963 AAGGTCAACTTGAGATGGTTAAAGCTGCAACCTGAGACGAATTTGCGGCTTCTGTTCTAC 1022  
Db 963 AAGGTCAACTTGAGATGGTTAAAGCTGCAACCTGAGACGAATTTGCGGCTTCTGTTCTAC 1022  
Qy 1023 CAGTTTCATAGATCCCATATTTGACTATCTGCTGCTCACTTTCACTTCTCTTCTGCAATAACA 1082  
Db 1023 CAGTTTCATAGATCCCATATTTGACTATCTGCTGCTCACTTTCACTTCTCTTCTGCAATAACA 1082  
Qy 1083 TCAAGCACCATACATATGCTTCAAGCAATTAATCTCAACATCCCAATCTTCAGTACCTTGA 1142  
Db 1083 TCAAGCACCATACATATGCTTCAAGCAATTAATCTCAACATCCCAATCTTCAGTACCTTGA 1142  
Qy 1143 TCCATAGCTTGGGGCTTCTTATACAGCAAGGCTCGATGAAACCAAGATGGAACGGA 1202  
Db 1143 TCCATAGCTTGGGGCTTCTTATACAGCAAGGCTCGATGAAACCAAGATGGAACGGA 1202  
Qy 1203 AAGATGTTCTTATAGAGCTTGGTCCATGGGCATATAGTTGAATTTACTTTCGACAGAGC 1262  
Db 1203 AAGATGTTCTTATAGAGCTTGGTCCATGGGCATATAGTTGAATTTACTTTCGACAGAGC 1262  
Qy 1263 AATTCTTGGAGATCTTCTCGAAGGCACAGCTTCTAGGAGTGGAACACCTCTTGTGCTC 1322  
Db 1263 AATTCTTGGAGATCTTCTCGAAGGCACAGCTTCTAGGAGTGGAACACCTCTTGTGCTC 1322  
Qy 1323 GGCAGGACTTTTGTCACTTGTGATGATACTCTGCTACCAATGTCAATCCAGACATCT 1382  
Db 1323 GGCAGGACTTTTGTCACTTGTGATGATACTCTGCTACCAATGTCAATCCAGACATCT 1382  
Qy 1383 TGATAATACCTGTTGGAATCTCTATGATGCGATATCGAAGTCACTACAAATGGTGAAC 1442  
Db 1383 TGATAATACCTGTTGGAATCTCTATGATGCGATATCGAAGTCACTACAAATGGTGAAC 1442  
Qy 1443 AACTGGCAAACTTAAGAAATGAGAGCTGTTGGAGTGTAGCAAGAGGTTTATTAGAA 1502  
Db 1443 AACTGGCAAACTTAAGAAATGAGAGCTGTTGGAGTGTAGCAAGAGGTTTATTAGAA 1502  
Qy 1503 TGTTACGAAAAAATATGGTTGTGTCGAGTGGATTTTGACAGCCATTTTCTTAAAGG 1562  
Db 1503 TGTTACGAAAAAATATGGTTGTGTCGAGTGGATTTTGACAGCCATTTTCTTAAAGG 1562  
Qy 1563 AATATTTAGAAAGCAAGTCAAGAACCGGTGTCTGCTTACTTTCCCTGGAGCAAGCGT 1622  
Db 1563 AATATTTAGAAAGCAAGTCAAGAACCGGTGTCTGCTTACTTTCCCTGGAGCAAGCGT 1622  
Qy 1623 TGTACCAGCTATCTTCTTCAAGACCCAGTGTGCTGATGAAGGTAGAGACAGCT 1682  
Db 1623 TGTACCAGCTATCTTCTTCAAGACCCAGTGTGCTGATGAAGGTAGAGACAGCT 1682  
Qy 1683 CCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGGAGGTTGATTGCAAAATC 1742  
Db 1683 CCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGGAGGTTGATTGCAAAATC 1742  
Qy 1743 TGCTGAGCATATTTCTATTCATCTGCTAGCAAGTCTGTGCGCATATGTCCACACATTTG 1802  
Db 1743 TGCTGAGCATATTTCTATTCATCTGCTAGCAAGTCTGTGCGCATATGTCCACACATTTG 1802  
Qy 1803 TGCTTGCCTGCTCTCTACAGACACAGCAGGGAATTTGATCTCTCCACATTTGGTGAAG 1862  
Db 1803 TGCTTGCCTGCTCTCTACAGACACAGCAGGGAATTTGATCTCTCCACATTTGGTGAAG 1862  
Qy 1863 ACTTCTTTGTGATGAAGGAGGAGTCTTGGCTCGTGAATTTGACCTGGGCTTCTCAGGAA 1922

Db 1863 ACTTCTTTGTGATGAAGGAGGAGTCTTGGCTCGTGAATTTTGACCTGGGCTTCTCAGGAA 1922  
Qy 1923 ATTCAGACATGTAGTAATGTCATGCCATACAGCTGCTGGGAAATTTGTGTACAAATCACCC 1982  
Db 1923 ATTCAGACATGTAGTAATGTCATGCCATACAGCTGCTGGGAAATTTGTGTACAAATCACCC 1982  
Qy 1983 ACATAGCAGGAACCGATGATGTTTATCACCCTCCAGCAACCTGCTCCATCAGTCTTCG 2042  
Db 1983 ACATAGCAGGAACCGATGATGTTTATCACCCTCCAGCAACCTGCTCCATCAGTCTTCG 2042  
Qy 2043 AACTCAACTTCTACAGCAATGGGTACTTTCATGCTTTTATCATGGAGGCAATCATAGCTT 2102  
Db 2043 AACTCAACTTCTACAGCAATGGGTACTTTCATGCTTTTATCATGGAGGCAATCATAGCTT 2102  
Qy 2103 GCAGCCTTTATGAGCTTCTGAACAGAGGGGACTGGGGGGTCCCACTAGCACCACCTA 2162  
Db 2103 GCAGCCTTTATGAGCTTCTGAACAGAGGGGACTGGGGGGTCCCACTAGCACCACCTA 2162  
Qy 2163 ACCTGATCAGCCAGGACGAGCTGGTGGGAAAGCGCCAGCCTGTGCTACCTTCTCTCCA 2222  
Db 2163 ACCTGATCAGCCAGGACGAGCTGGTGGGAAAGCGCCAGCCTGTGCTACCTTCTCTCCA 2222  
Qy 2223 ATGAAGGCAACCTCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAAACAGTAG 2282  
Db 2223 ATGAAGGCAACCTCTCACTGCTTGCAGACATTTTACCAAGTCTGCCATGAAACAGTAG 2282  
Qy 2283 GAAAGTTTATCAGTATGTCATTTACAGTGGCAGACGATGATGATGATATCA 2342  
Db 2283 GAAAGTTTATCAGTATGTCATTTACAGTGGCAGACGATGATGATGATATCA 2342  
Qy 2343 GTCCTAGTCTGTGAGCAGTGGGCAAGAGCTTCCAGAACCTTTGTCTTGGAGAA 2402  
Db 2343 GTCCTAGTCTGTGAGCAGTGGGCAAGAGCTTCCAGAACCTTTGTCTTGGAGAA 2402  
Qy 2403 GTGATGAAGAAATGAGACAGTGAATTTGGGAGGAAACAGGAGATGCTACCTGAAGG 2462  
Db 2403 GTGATGAAGAAATGAGACAGTGAATTTGGGAGGAAACAGGAGATGCTACCTGAAGG 2462  
Qy 2463 TGAGCCAAATCCAGGAGCAGCAGTGTATCACCCTTTTACAGAGACTTCTTGGGCTT 2522  
Db 2463 TGAGCCAAATCCAGGAGCAGCAGTGTATCACCCTTTTACAGAGACTTCTTGGGCTT 2522  
Qy 2523 TGCTGAGGCTTACAGCTCTGCTGCAATCTTGTTCACAACTTTCAGTGTCTGTTCCAG 2582  
Db 2523 TGCTGAGGCTTACAGCTCTGCTGCAATCTTGTTCACAACTTTCAGTGTCTGTTCCAG 2582  
Qy 2583 AACCTGAGTATCTGCAAAAGTTGCAAAATACCTAATAACCAAGAAAGAAATGTTG 2642  
Db 2583 AACCTGAGTATCTGCAAAAGTTGCAAAATACCTAATAACCAAGAAAGAAATGTTG 2642  
Qy 2643 CAGTATATGCTGAGAGTGCACATATTTGTTGAGAAATGCTGTGAAATGTTTAAAGG 2702  
Db 2643 CAGTATATGCTGAGAGTGCACATATTTGTTGAGAAATGCTGTGAAATGTTTAAAGG 2702  
Qy 2703 ATATTGGGCTTTTCAAGGAGACCAAAAGAGAGTGTCTGTTTGTAGAACTTGAGCAGCA 2762  
Db 2703 ATATTGGGCTTTTCAAGGAGACCAAAAGAGAGTGTCTGTTTGTAGAACTTGAGCAGCA 2762  
Qy 2763 CTTTCTACCTCAATGCAACCGAACAAAATCTTCTAGAATATATCTGAGTTTGTGGTGC 2822  
Db 2763 CTTTCTACCTCAATGCAACCGAACAAAATCTTCTAGAATATATCTGAGTTTGTGGTGC 2822  
Qy 2823 TGTAGTAACTGCTGGCAGTGTGGCAATGAAAGTTCATGAGATGATGATGATGATGATGAT 2882  
Db 2823 TGTAGTAACTGCTGGCAGTGTGGCAATGAAAGTTCATGAGATGATGATGATGATGATGAT 2882  
Qy 2883 CCAGCTTCTGGCTCAAGAGTGTGGCAATGAAAGTTCATGAGATGATGATGATGATGATGAT 2942  
Db 2883 CCAGCTTCTGGCTCAAGAGTGTGGCAATGAAAGTTCATGAGATGATGATGATGATGATGAT 2942  
Qy 2943 AAGTGAATCTCTGGAGACAGTGTCTTCCATCGATCTGAGATCTTCTCCAGCTT 3002  
Db 2943 AAGTGAATCTCTGGAGACAGTGTCTTCCATCGATCTGAGATCTTCTCCAGCTT 3002

```
QY      3003 T 3003
Db      3001 T 3001

RESULT 3
US-10-288-252-26
; Sequence 26, Application US/10288252
; Publication No. US20030143686A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAL, Preeti G.
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: BURFORD, Neil
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: YAO, Monique G.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAPALIA, April J.A.
; APPLICANT: LU, Yan
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: SANJANWALA, Madhu S.
; APPLICANT: DING, Li
; TITLE OF INVENTION: TRANSFERASES
; FILE REFERENCE: PI-0241 USA
; CURRENT APPLICATION NUMBER: US/10/288,252
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT US01/30424
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/252,819
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/249,639
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/247,931
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/246,001
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/244,025
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/238,481
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,523
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 26
; LENGTH: 3210
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030143686A1 1517230CB1
US-10-288-252-26

Query Match      98.5%; Score 2958; DB 13; Length 3210;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2983; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY      12 GCTGCTGGGGGGGACTCTTTCTGAGGTTACTGTGGAGCACCACCAAGTCTGTGAGCCCTCT 71
Db      9 GCTGCTGGGGGGGACTCTTTCTGAGGTTACTGTGGAGCACCACCAAGTCTGTGAGCCCTCT 68

QY      72 GGCCGTGCAAAACAGGACCCACGAGGAACAGACCTTGTCTTAATTCACCAACAGCCCTGGGAC 131
Db      69 GGCCGTGCAAAACAGGACCCACGAGGAACAGACCTTGTCTTAATTCACCAACAGCCCTGGGAC 128

QY      132 TGTCTTCTCCAGAGTCTCATCAGCTTGTGCTTAATCGACTGATGGAATAATTTCTTCAA 191
Db      129 TGTCTTCTCCAGAGTCTCATCAGCTTGTGCTTAATCGACTGATGGAATAATTTCTTCAA 188

192 CACCACCAAGTCAAGGATACAGGACGAGCGGCTCCCTGTTGTATGACATTTCTGCACC 251
189 CACCACCAAGTCAAGGATACAGGACGAGCGGCTCCCTGTTGTATGACATTTCTGCACC 248
252 CGAAACTGATAGCTGAGTCTGAAAGTTTATGTTATGAAACAGAGAAGAACTTTTCATCCAG 311
249 CGAAACTGATAGCTGAGTCTGAAAGTTTATGTTATGAAACAGAGAAGAACTTTTCATCCAG 308
312 CACATGATTTGGGAATTACACTTTGTGACATGGAATGATCTGCACTGACCCCTTGGTACAA 371
309 CACATGATTTGGGAATTACACTTTGTGACATGGAATGATCTGCACTGACCCCTTGGTACAA 368
372 TAGATGTTTCTTATCTGCCACATTCATCAGAAATACAGTGTGGTGGTGAAGTGAAGCACAA 431
369 TAGATGTTTCTTATCTGCCACATTCATCAGAAATACAGTGTGGTGGTGAAGTGAAGCACAA 428
432 GTGAGGAATGGGTGAGTGTGGCTTTAGACCCACCGTCTTCAGATCTGCAAACTTTAAAT 491
429 GTGAGGAATGGGTGAGTGTGGCTTTAGACCCACCGTCTTCAGATCTGCAAACTTTAAAT 488
492 GGAAGAAAGCCTAATGAGTCGGAAGAGCCATTTGTTGGAAGATGTTGTTACTCTGCA 551
489 GGAAGAAAGCCTAATGAGTCGGAAGAGCCATTTGTTGGAAGATGTTGTTACTCTGCA 548
552 CTCCCAGAGCTGGGACAAATTTTCAACCCAGTATCCCGTCTTTGGGTTTGGGGAATG 611
549 CTCCCAGAGCTGGGACAAATTTTCAACCCAGTATCCCGTCTTTGGGTTTGGGGAATG 608
612 TTATTTATATCAATGAAACTCACAACAAGACCCGGATGGCTTGAAGACGCTTTCTT 671
609 TTATTTATATCAATGAAACTCACAACAAGACCCGGATGGCTTGAAGACGCTTTCTT 668
672 ACGTCTTTTATCAAGAGCGAGATGTCATAGGGCATGTTTCCACCAATGTGACTG 731
669 ACGTCTTTTATCAAGAGCGAGATGTCATAGGGCATGTTTCCACCAATGTGACTG 728
732 AAAATGTGCTGAACAGCAGTAGAGTACAAGAGGCAATTTGCAAGAGTGGCTCTCAATTA 791
729 AAAATGTGCTGAACAGCAGTAGAGTACAAGAGGCAATTTGCAAGAGTGGCTCTCAATTA 788
792 ACCGTGATGTTTCTGCCAGCAGCAATCAAAGCCGTTAAACAAAGTGAAGAAAGTAA 851
789 ACCGTGATGTTTCTGCCAGCAGCAATCAAAGCCGTTAAACAAAGTGAAGAAAGTAA 848
852 AAAGGATTTCAAGAAATGGTTGCCACTGTCTCACCGGCAATCATCAGACTGCTGGGT 911
849 AAAGGATTTCTCAAGAAATGGTTGCCACTGTCTCACCGGCAATCATCAGACTGCTGGGT 908
912 GGGTGTGCTAAAACTGTTCAACAGCTTCTTTTGGAAACATTTCAAAATTCACAAAGTCAAC 971
909 GGGTGTGCTAAAACTGTTCAACAGCTTCTTTTGGAAACATTTCAAAATTCACAAAGTCAAC 968
972 TTGAGATGGTTAAAGCTGCAACTGAGACGAATTTGGCGCTTCTGTTTCTACAGTTCATA 1031
969 TTGAGATGGTTAAAGCTGCAACTGAGACGAATTTGGCGCTTCTGTTTCTACAGTTCATA 1028
1032 GATCCCATTTGACTATCTGCTGCTCACTTTCTCTTCTGCCATAACATCAAGAC 1091
1029 GATCCCATTTGACTATCTGCTGCTCACTTTCTCTTCTGCCATAACATCAAGAC 1088
1092 CATACATTTGCTTCAAGCAATAATCTCAACATCCCAATCTTCTGAGTACCTTGATCCATAGC 1151
1089 CATACATTTGCTTCAAGCAATAATCTCAACATCCCAATCTTCTGAGTACCTTGATCCATAGC 1148
1152 TTGGGGCTTTCTTCATACGACGAAGGCTCGATGAAACACAGATGCGGAAAGATGTTTC 1211
1149 TTGGGGCTTTCTTCATACGACGAAGGCTCGATGAAACACAGATGCGGAAAGATGTTTC 1208
1212 TCTATAGAGCTTTGCTCCATGGGCAATATAGTTGAATTTACTTCGACAGAGCAATTTCTTGG 1271
1209 TCTATAGAGCTTTGCTCCATGGGCAATATAGTTGAATTTACTTCGACAGAGCAATTTCTTGG 1268
```

QY 1272 AGATCTTCTGGAAGGACACAGTCTTAGAGAGTGGAAAAAAGCTCTTGTGCTCGGCGAGGAC 1331  
Db 1269 AGATCTTCTGGAAGGACACAGTCTTAGAGAGTGGAAAAAAGCTCTTGTGCTCGGCGAGGAC 1328  
QY 1332 TTTTGTCAAGTGTGTAGATCTCTGTCTACCAATGTCATCCAGACATCTTGTATATAC 1391  
Db 1329 TTTTGTCAAGTGTGTAGATCTCTGTCTACCAATGTCATCCAGACATCTTGTATATAC 1388  
QY 1392 CTGTGTGAATCTCTATGATCGCATATCGAAGGTCACTACAATGGTGAACAACCTGGGCA 1451  
Db 1389 CTGTGTGAATCTCTATGATCGCATATCGAAGGTCACTACAATGGTGAACAACCTGGGCA 1448  
QY 1452 AACCTAAGAAAGATGAGAGCCTGTGGAGTGTAGCAAGAGGTGTATTTAGAAATGTACGAA 1511  
Db 1449 AACCTAAGAAAGATGAGAGCCTGTGGAGTGTAGCAAGAGGTGTATTTAGAAATGTACGAA 1508  
QY 1512 AAAAATATGTTGTGTCGAGTGGATTTTGCACAGCCATTTTCTTTAAAGGAATATTTAG 1571  
Db 1509 AAAAATATGTTGTGTCGAGTGGATTTTGCACAGCCATTTTCTTTAAAGGAATATTTAG 1568  
QY 1572 AAAGCCAAAGTCAGAAACCGGTGTCTCTACTTTTCCCTGGAGCAAGCGTGTGTACCGAG 1631  
Db 1569 AAAGCCAAAGTCAGAAACCGGTGTCTCTACTTTTCCCTGGAGCAAGCGTGTGTACCGAG 1628  
QY 1632 CTATACCTTCTTCAAGACCCAGTGTCTGTCTGATGAAGGTGAGACACAGTCCATTAATG 1691  
Db 1629 CTATACCTTCTTCAAGACCCAGTGTCTGTCTGATGAAGGTGAGACACAGTCCATTAATG 1688  
QY 1692 AGTCCAGAAATGCAACAGATGAATCCCTAGCAAGGAGGTGTGATTCGAAATCTGGCTGAGC 1751  
Db 1689 AGTCCAGAAATGCAACAGATGAATCCCTAGCAAGGAGGTGTGATTCGAAATCTGGCTGAGC 1748  
QY 1752 ATATTCTTATCTAGTGTAGCAAGTCTGTGCTCAATTTATGTCACACACATTTGCGCTTGCC 1811  
Db 1749 ATATTCTTATCTAGTGTAGCAAGTCTGTGCTCAATTTATGTCACACACATTTGCGCTTGCC 1808  
QY 1812 TGCTCTCTACAGACACAGCAGGGAATTTGATCTCTCCATTTGGTGGAGACTTCTTTG 1871  
Db 1809 TGCTCTCTACAGACACAGCAGGGAATTTGATCTCTCCATTTGGTGGAGACTTCTTTG 1868  
QY 1872 TGATGAAGGAAGTCTGGCTCGTGATTTTGAACCTGGGGTCTCAGGAAATTCAGAG 1931  
Db 1869 TGATGAAGGAAGTCTGGCTCGTGATTTTGAACCTGGGGTCTCAGGAAATTCAGAG 1928  
QY 1932 ATCTAGTATGATGCATACAGTCTGTGGAAATTTGTGTCACAAATCCACCACACTAGCA 1991  
Db 1929 ATCTAGTATGATGCATACAGTCTGTGGAAATTTGTGTCACAAATCCACCACACTAGCA 1988  
QY 1992 GGAACGATGAGTTTTTTATCACCCCGACGACAACTGTCCCATCAGTCTTCGAACTCAACT 2051  
Db 1989 GGAACGATGAGTTTTTTATCACCCCGACGACAACTGTCCCATCAGTCTTCGAACTCAACT 2048  
QY 2052 TCTACAGCAATGGGGTACTTTCATGCTTTATCATATGAGGCGCATCATAGCTTGCAGCCTTT 2111  
Db 2049 TCTACAGCAATGGGGTACTTTCATGCTTTATCATATGAGGCGCATCATAGCTTGCAGCCTTT 2108  
QY 2112 ATTCAGTCTTGAACAGAGGGGACTGGGGGTCCCATAGCACCCACCTAACCTCATCA 2171  
Db 2109 ATTCAGTCTTGAACAGAGGGGACTGGGGGTCCCATAGCACCCACCTAACCTCATCA 2168  
QY 2172 GCCAGGAGCAGCTGGTGGGAAAGCGGCGCAGCCTGTGCTACTTCTTCCAATGAAGGCA 2231  
Db 2169 GCCAGGAGCAGCTGGTGGGAAAGCGGCGCAGCCTGTGCTACTTCTTCCAATGAAGGCA 2228  
QY 2232 CCATCTCAGTCCCTTGCAGACATTTTATCAAGTGTGCCCATGAAACAGTAGGAAAGTTTA 2291  
Db 2229 CCATCTCAGTCCCTTGCAGACATTTTATCAAGTGTGCCCATGAAACAGTAGGAAAGTTTA 2288  
QY 2292 TCCAGTATGCAATTTCTTACAGTGGCAGAGCAGCATGACACAGGAGATATCAGTCCCTAGTC 2351  
Db 2289 TCCAGTATGCAATTTCTTACAGTGGCAGAGCAGCATGACACAGGAGATATCAGTCCCTAGTC 2348  
QY 2352 TTGCTGAGCAGCAGTGGGACAAAGAGCTTCCAGAACTTTTGTCTTGGAGAAAGTGATGAAG 2411

Db 2349 TTGCTGAGCAGCAGTGGGACAAAGAGCTTCTTGAACCTTTGTCTTGGAGAAAGTGATGAAG 2408  
QY 2412 AAGATCAACACAGTGTACTTTTGGGGAGGAACAGCAGATTTGCTACCTGGAAGGTGAGCAAT 2471  
Db 2409 AAGATCAACACAGTGTACTTTTGGGGAGGAACAGCAGATTTGCTACCTGGAAGGTGAGCAAT 2468  
QY 2472 CCAAGGAGCACCAGCAGTGTATACCTTCTTACAGAGACTCCTTGGGCCCTTTGCTGGAGG 2531  
Db 2469 CCAAGGAGCACCAGCAGTGTATACCTTCTTACAGAGACTCCTTGGGCCCTTTGCTGGAGG 2528  
QY 2532 CTTACAGCTCTGCTGCTCATCTTTGTTTCAAACTTTAGTGGTCTGTTCAGAAACCTGAGT 2591  
Db 2529 CTTACAGCTCTGCTGCTCATCTTTGTTTCAAACTTTAGTGGTCTGTTCAGAAACCTGAGT 2588  
QY 2592 ATCTGCAAAAGTTGCACAAATACCTTAATAACAGAAACAGAAAGAAATGTTGCAAGTATATG 2651  
Db 2589 ATCTGCAAAAGTTGCACAAATACCTTAATAACAGAAACAGAAAGAAATGTTGCAAGTATATG 2648  
QY 2652 CTGAGAGTCCCATATATTTCTTGTGAAGAAATGCTGTGAAATGTTTAAAGGATATTTGGGG 2711  
Db 2649 CTGAGAGTCCCATATATTTCTTGTGAAGAAATGCTGTGAAATGTTTAAAGGATATTTGGGG 2708  
QY 2712 TTTTCAAGGAGACCAAAACAAAGAGAGTGTCTGTTTTAGAACTGAGCAGACACTTTTCTAC 2771  
Db 2709 TTTTCAAGGAGACCAAAACAAAGAGAGTGTCTGTTTTAGAACTGAGCAGACACTTTTCTAC 2768  
QY 2772 CTCATATGCAACCGACCAAAACCTTCTAGAAATATATCTGAGTGTGTTGGTGTCTAGGTA 2831  
Db 2769 CTCATATGCAACCGACCAAAACCTTCTAGAAATATATCTGAGTGTGTTGGTGTCTAGGTA 2828  
QY 2832 CGTGTGGCACTGTGCGCAAAATGAAGTCAATGAGATGAGTTCCTTTGATGATACCAGCTTCT 2891  
Db 2829 CGTGTGGCACTGTGCGCAAAATGAAGTCAATGAGATGAGTTCCTTTGATGATACCAGCTTCT 2888  
QY 2892 GGTCTAAGAGTTTGAAGTGCCTTCCAGGGGTGAGCCTGCGCCTGTCNCGAAAGTGAATCT 2951  
Db 2889 GGTCTAAGAG-ITGAAGGTGCGGTGCGCA-GGGTCAAGGCTGCGCCTGTCNCGAAAGTGAATCT 2946  
QY 2952 CTTGGAAGCAAGTGCCTTCTNCCCTCAATGGAATCTGAGATCTTCCAGCTTT 3003  
Db 2947 CTTGGAAGCAAGTGCCTTCTNCCCTCAATGGAATCTGATCTTCCAGCTTCT 2998

## RESULT 4

US-09-935-290-3  
; Sequence 3, Application US/09935290  
; Publication No. US2003004948A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller-Libermann, Rosana  
; TITLE OF INVENTION: 56919, A NOVEL HUMAN ACYLTRANSFERASE AND USES THEREOF  
; FILE REFERENCE: MNI-186  
; CURRENT APPLICATION NUMBER: US/09/935,290  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/226,509  
; PRIOR FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2484  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-935-290-3

Query Match 82.7%; Score 2484; DB 11; Length 2484;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 341 ATGGATGAATCTGCACTGACCTTGGTACAATAGATGTTTCTTATCTGCCACATTCATCA 400  
Db 1 ATGGATGAATCTGCACTGACCTTGGTACAATAGATGTTTCTTATCTGCCACATTCATCA 60  
QY 401 GAATACAGTGTGTTGTCGATGTAAGCACACAGTGAGAAATGGGGTGAAGTGTGCTTTAGA 460

Db 61 |||||GAATACAGTGTGGTCGATGTAAGCACACAGTGGGAATGGGTGAGTGGCTTTTAGA 120  
Qy 461 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAAATCAGTTCGGAAGG 520  
Db 121 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAAATCAGTTCGGAAGG 180  
Qy 521 CCATTTGTTGGAAGATGTTGTTATCTCTGCACTCCCCAGAGCTGGGACAAATTTTCAAC 580  
Db 181 CCATTTGTTGGAAGATGTTGTTATCTCTGCACTCCCCAGAGCTGGGACAAATTTTCAAC 240  
Qy 581 CCCAGTATCCCGTCTTTGGGTTCGGGAATGTTATTTATATCAATGAACCTCACACAGA 640  
Db 241 CCCAGTATCCCGTCTTTGGGTTCGGGAATGTTATTTATATCAATGAACCTCACACAGA 300  
Qy 641 CACCGCGATGGCTTTCGAAGACGCCCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTG 700  
Db 301 CACCGCGATGGCTTTCGAAGACGCCCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTG 360  
Qy 701 CATTAAGGCGATGTTTGGCCACCAATGTGATGTAAGAAATGTCTGAAACAGCAGTAGAGTCAA 760  
Db 361 CATTAAGGCGATGTTTGGCCACCAATGTGATGTAAGAAATGTCTGAAACAGCAGTAGAGTCAA 420  
Qy 761 GAGGCAATTCAGAGTGGCTGCTGAATTAACCTGATGGTCTGCCCCAGCAGCAATCA 820  
Db 421 GAGGCAATTCAGAGTGGCTGCTGAATTAACCTGATGGTCTGCCCCAGCAGCAATCA 480  
Qy 821 AAAGCCGTTAAAGATGAAAAAGAAAGCTTAAAGGAATCTTCAAGAAATGGTTGCCACT 880  
Db 481 AAAGCCGTTAAAGATGAAAAAGAAAGCTTAAAGGAATCTTCAAGAAATGGTTGCCACT 540  
Qy 881 GTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGCTAAACCTGTTCAACAGCTTC 940  
Db 541 GTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGCTAAACCTGTTCAACAGCTTC 600  
Qy 941 TTTTGGACATTCAAATTCACAAAGTCAACTGAGATGGTTAAAGCTGCAACTGAGAGC 1000  
Db 601 TTTTGGACATTCAAATTCACAAAGTCAACTGAGATGGTTAAAGCTGCAACTGAGAGC 660  
Qy 1001 AATTGGCCGTTCTGTTTCTACCAAGTTCATAGATCCCATATTTGACTATCTGCTCACT 1060  
Db 661 AATTGGCCGTTCTGTTTCTACCAAGTTCATAGATCCCATATTTGACTATCTGCTCACT 720  
Qy 1061 TTCAATCTCTTCTGCAATACATCAAGCACCACATACATTTGCTTCAGGCAATAATCTCAAC 1120  
Db 721 TTCAATCTCTTCTGCAATACATCAAGCACCACATACATTTGCTTCAGGCAATAATCTCAAC 780  
Qy 1121 ATCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTCATACGACGAGGCTC 1180  
Db 781 ATCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTCATACGACGAGGCTC 840  
Qy 1181 GATGAACACACAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1240  
Db 841 GATGAACACACAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Qy 1241 GTTGAATTTACTTCGACACAGCAATTCITGGAGATCTTCTGGAAGGCACACGTTCTAGG 1300  
Db 901 GTTGAATTTACTTCGACACAGCAATTCITGGAGATCTTCTGGAAGGCACACGTTCTAGG 960  
Qy 1301 AGTGAAACAACTCTTGTGCTCGGCAGGACTTTTGTGTCAGTGTGGTAGATCTGTGCT 1360  
Db 961 AGTGAAACAACTCTTGTGCTCGGCAGGACTTTTGTGTCAGTGTGGTAGATCTGTGCT 1020  
Qy 1361 ACCAATGTCATCCAGACATCTGATATATCTGTTGGAATCTCTATAGAGCTTCATGCAATATC 1420  
Db 1021 ACCAATGTCATCCAGACATCTGATATATCTGTTGGAATCTCTATAGAGCTTCATGCAATATC 1080  
Qy 1421 GAAGTCACTACAATGTTGAACAACTGGGCAAACTTAAGAAATGAGAGCTGTGGAGT 1480  
Db 1081 GAAGTCACTACAATGTTGAACAACTGGGCAAACTTAAGAAATGAGAGCTGTGGAGT 1140  
Qy 1481 GTAGCAAGAGTGTATTAGAAATGTTACGAAAAAACTATAGTTGTGTCGAGTGGATTTT 1540

Db 1141 GTACCAAGAGGTGTTATTAGAAATGTTACGAAAAAACTATGGTGTGTCGAGTGGATTTT 1200  
Qy 1541 GCACAGCCATTTTCTCTTAAAGGAATATTATGAAGCCAAAGCTCAGAAACCGGTGTCTGCT 1600  
Db 1201 GCACAGCCATTTTCTCTTAAAGGAATATTATGAAGCCAAAGCTCAGAAACCGGTGTCTGCT 1260  
Qy 1601 CTACTTTCCCTGGAGCAAGCGTGTACCACTATATCTTCCCTCAAGACCCAGTGTGCT 1660  
Db 1261 CTACTTTCCCTGGAGCAAGCGTGTACCACTATATCTTCCCTCAAGACCCAGTGTGCT 1320  
Qy 1661 GCTGATCAAGGTAGAGACAGTCCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1720  
Db 1321 GCTGATCAAGGTAGAGACAGTCCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Qy 1721 CGAAGGAGGTGATTGTCAAATCTGGCTGAGCATATCTATTCTACTGTAGCAAGTCCCTGT 1780  
Db 1381 CGAAGGAGGTGATTGTCAAATCTGGCTGAGCATATCTATTCTACTGTAGCAAGTCCCTGT 1440  
Qy 1781 GCCATTATGTCACACATTTGTGGTGTGCTGCTCTCTACAGACACAGCAGGGAAT 1840  
Db 1441 GCCATTATGTCACACATTTGTGGTGTGCTGCTCTCTACAGACACAGCAGGGAAT 1500  
Qy 1841 GATCTCTCCATTTGTCGAAAGCTTCTTTGTGATGAAGAGGAAGTCTTGGCTCGTGT 1900  
Db 1501 GATCTCTCCATTTGTCGAAAGCTTCTTTGTGATGAAGAGGAAGTCTTGGCTCGTGT 1560  
Qy 1901 TTTGACCTGGGTCTCAGGAAATTCAGAAAGTGTAGTAAATGATGTCATGTCATACAGTGTG 1960  
Db 1561 TTTGACCTGGGTCTCAGGAAATTCAGAAAGTGTAGTAAATGATGTCATGTCATACAGTGTG 1620  
Qy 1961 GGAATTTGTGTCAATCAATCCCACTAGCAGGAAACGATGATTTTTTATCACCCCCAGC 2020  
Db 1621 GGAATTTGTGTCAATCAATCCCACTAGCAGGAAACGATGATTTTTTATCACCCCCAGC 1680  
Qy 2021 ACAATGTCCCATCAGTCTTCGAAACTCACTTCAGAGCAATGGGGTACTTCTATGTCTTT 2080  
Db 1681 ACAATGTCCCATCAGTCTTCGAAACTCACTTCAGAGCAATGGGGTACTTCTATGTCTTT 1740  
Qy 2081 ATCATGAGGCGCATCATAGCTTGCAGCTTTATGTCAGTTCCTGAACAGAGGGGACTGGG 2140  
Db 1741 ATCATGAGGCGCATCATAGCTTGCAGCTTTATGTCAGTTCCTGAACAGAGGGGACTGGG 1800  
Qy 2141 GGTCCCACTAGCAGCCCACTTAACCTGATCAGCAGGAGCAGCTGTGCGGAAGCGGCC 2200  
Db 1801 GGTCCCACTAGCAGCCCACTTAACCTGATCAGCAGGAGCAGCTGTGCGGAAGCGGCC 1860  
Qy 2201 AGCTGTGCTACCTTCTCTCAATGAAGGACCATCTCACTGCTTGCAGACATTTTAC 2260  
Db 1861 AGCTGTGCTACCTTCTCTCAATGAAGGACCATCTCACTGCTTGCAGACATTTTAC 1920  
Qy 2261 CAAGTCTGCCATGAAACAGTAGGAAGTATTATCCAGTATGCAATTTTACAGTGGCAGAG 2320  
Db 1921 CAAGTCTGCCATGAAACAGTAGGAAGTATTATCCAGTATGCAATTTTACAGTGGCAGAG 1980  
Qy 2321 CACGATCAGCAGGAAGATATCAGTCTTAGTCTTCTGAGCAGCAGTGGGCAAGAGCTT 2380  
Db 1981 CACGATCAGCAGGAAGATATCAGTCTTAGTCTTCTGAGCAGCAGTGGGCAAGAGCTT 2040  
Qy 2381 CCAGAACTTTTGTCTTGGAGAGTGTATGAAGAGATGAAGAGTGAAGAGTGAAGAGTGAAG 2440  
Db 2041 CCAGAACTTTTGTCTTGGAGAGTGTATGAAGAGATGAAGAGTGAAGAGTGAAGAGTGAAG 2100  
Qy 2441 CAGCGAGATTTGCTTACCTGAGGCTTACCAATCCAGGAGCAGCAGAGTATTACCTTC 2500  
Db 2101 CAGCGAGATTTGCTTACCTGAGGCTTACCAATCCAGGAGCAGCAGAGTATTACCTTC 2160  
Qy 2501 TTACAGAGATCTCTTGGGCTTTCGAGGCTTACAGCTCTGCTGCTCATCTTTGTTTAC 2560  
Db 2161 TTACAGAGATCTCTTGGGCTTTCGAGGCTTACAGCTCTGCTGCTCATCTTTGTTTAC 2220  
Qy 2561 AACTTCAGTGGTCTGTTTCAGAACTGAGTATCTGCAAAAGTTGCAAAATACCTAATA 2620  
Db 2221 AACTTCAGTGGTCTGTTTCAGAACTGAGTATCTGCAAAAGTTGCAAAATACCTAATA 2280

QY 2621 ACCAGAACAGAAAGAAATGTTGAGATATGCTGAGAGTGCCACATATGCTTTGTGAAG 2680  
DB |||||  
QY 2281 ACCAGAACAGAAAGAAATGTTGAGATATGCTGAGAGTGCCACATATGCTTTGTGAAG 2340  
DB |||||  
QY 2681 AATGCTGTGAAATGTTTAAGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTG 2740  
DB |||||  
QY 2341 AATGCTGTGAAATGTTTAAGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTG 2400  
DB |||||  
QY 2741 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAACAACTTCTAGAA 2800  
DB |||||  
QY 2401 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAACAACTTCTAGAA 2460  
DB |||||  
QY 2801 TATATTTCTGAGTTTGTGGTGCTG 2824  
DB |||||  
QY 2461 TATATTTCTGAGTTTGTGGTGCTG 2484  
DB |||||

## RESULT 5

US-09-917-800A-1648  
; Sequence 1648, Application US/09917800A  
; Patent No. US20020119462A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendrick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Castle, Arthur  
; APPLICANT: Elashoff, Michael  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCE: 44921-5038-US  
; CURRENT APPLICATION NUMBER: US/09/917,800A  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,040  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,880  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: US 60/290,029  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US 60/290,645  
; PRIOR FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: US 60/292,336  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/295,798  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/297,457  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,884  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,459  
; PRIOR FILING DATE: 2001-07-09  
; NUMBER OF SEQ ID NOS: 1740  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1648  
; LENGTH: 2646  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM\_017274  
US-09-917-800A-1648

Query Match 63.1%; Score 1896.2; DB 10; Length 2646;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 2126; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

QY 319 TTGGGAATTTACACTTTGTGACATGGATGAATCTGCACCTGACCTTGGTGAATAGATGT 378  
DB |||||  
QY 138 TTGGGACCTCCACGTTGAGTCATGGAGAGTCTTCAGTGACAAATGGCACAATAGACGT 197  
DB |||||  
QY 379 TTCTTATCTGCCACATTCATCAGATACAGTGTGGTCGATGTAAGCACAAGAGTGAAG 438  
DB |||||  
QY 198 TTCTTATCTGCCAAATTCATCAGAAATACAGCCTTGGCCGATGTAACACACGAATGAG 257  
DB |||||

QY 439 ATGGGCTGAGTGTGGCTTTAGACCCACCGCTCTTCAGATCTGCAACTTTAAATGGAAGA 498  
DB |||||  
QY 258 CTGGGTTGACTGTGGCTTCAAACTTCTTCAGATCCGCAACGCTGAAATGGAAGGA 317  
DB |||||  
QY 499 AAGCCTAATGAGTCGAAAAAGGCCATTTGTTGGAAGATGTTGTACTCTGCACTCCCCA 558  
DB |||||  
QY 318 GAGCCTCATGAGCCGGAAGAGGCCCTTCTGTTGGAAGGTGTGCTATTTCATGCAACGCTCA 377  
DB |||||  
QY 559 GAGCTGGGACAAATTTTCAACCCAGTATCCGCTCTTTGGGTTTTCGCGAATGTTTATTA 618  
DB |||||  
QY 378 GAGCTGGGAAGGTTTTTCAACCCAGTATCCATCTCTGGGTTTTCGCGAATGTTTATTA 437  
DB |||||  
QY 619 TATCAATGAAACTCACACAAGACACCGCGGATGGCTTGAAGAAGCGCTTTCTTACGTTCT 678  
DB |||||  
QY 438 TATCAATGAACTCACACAAGGACCGAGGATGGCTGGCAAGAGCGGCTTTCTTACATCCT 497  
DB |||||  
QY 679 TTTTATTCAAGAGCGAGATGTGCATGAAGGCGATGTTTGCACCAATGTGACTGAAATGT 738  
DB |||||  
QY 498 TTTTGTTCAGAGCGCGATGTCCACAAGGGCGATGTTTGGCCACCGATCATCTGACAATGT 557  
DB |||||  
QY 739 GCTGAACAGCAGTAGATGACAGGCAATTTGCAAGAGTGCGCTGCTGAAATTTAAACCTGA 798  
DB |||||  
QY 558 ACTGAATAGCAGAGATCCAAAGAGCAATTTGCTGAAGTGGCTGCAAGATTTGAACCCGGA 617  
DB |||||  
QY 799 TGGTTCTGCCAGCAGCAATCAAAAGCCGTTTAAACAAAGTGAAGAAAGAGCTTAAAGGAT 858  
DB |||||  
QY 618 TGGATCTGCCAGCAGCAGTCCAAAGCCATCCAGAAAGTGAAGAAAGCCAGGAAGAT 677  
DB |||||  
QY 859 TCTTCAAGAAATGTTGGCCACTGTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCT 918  
DB |||||  
QY 678 CTTCCAGGAAATGTTGTGTACAGTCTCCCGGGATGATCAGGCTGACTGGCTGGGTGTT 737  
DB |||||  
QY 919 GCTTAAACTGTTCAACAGCTTCTTTTGGAAACATTTCAAAATTCACAAAGTCAACTTCAGAT 978  
DB |||||  
QY 738 ACTAAAGCTCTTCAACAGCTTCTTCTGGAACATTTAGATTCAGAGGTCAACTTGAGAT 797  
DB |||||  
QY 979 GGTAAAGCTGCAACTGAGACGAAATTTGCGGCTTCTGTTTCTACCAAGTTTCATAGATCCCA 1038  
DB |||||  
QY 798 GGTAAAGCTGCAACTGAGACGAAATCTGCGGCTCTGTTTCTGCGGTGACAGATCCCA 857  
DB |||||  
QY 1039 TATGACTATCTGCTGCTCACTTTTCACTTCTTCTGCGCAATACATCAAAAGCACCACAT 1098  
DB |||||  
QY 858 CATCGACTACTGCTGCTCACTTCTCTCTCTTCTGSCCAACATCAAAAGCTCCATACAT 917  
DB |||||  
QY 1099 TSCCTTCAGGCAATTAATCTCAACATCCCAATCTTCAGTACCTTTCATCAAGCTTCGCGG 1158  
DB |||||  
QY 918 CGCTTCGGGCAACACCTCAACATCCCACTCTTCAGTACCTTTCATCAAGCTTCGCGG 977  
DB |||||  
QY 1159 CTTCTTTCATACGACGAGGCTCGATGAACACACAGATGGACGAAAGATGTTTCTTATAG 1218  
DB |||||  
QY 978 CTTTTCATAGACGAGGCTTTCAGCAACTCCAGATGGACGCAACATCTGTACAG 1037  
DB |||||  
QY 1219 AGCTTTGCTCCATGGGCATATAGTTGAATTAATCTTCAGACGAGCAATTTCTTGAGATCTT 1278  
DB |||||  
QY 1038 AGCGTTGCTCCATGGGCATATAGTTGAATCTCTCCAGACGAGCAGTTCCTTGAGATCTT 1097  
DB |||||  
QY 1279 CTTGGAAGGCACACGTTCTAGGAGTGGAACCTCTTCGCTCGGCGAGGACTTTTGTTC 1338  
DB |||||  
QY 1098 CTTGGAAGGCACCCGCTCCCGAGTGGCAAGACCTCTGTCGCGGCGGCTCTCTGTC 1157  
DB |||||  
QY 1339 AGTTGCTGATGATCTCTGCTACCAATGTCATCCAGACATCTTGTATATACCTGTTGG 1398  
DB |||||  
QY 1158 AGTTGCTGATGATCTCTGCTCATCCACACCATCCCTGACATCTCTGTCATCCCTGTTGG 1217  
DB |||||  
QY 1399 AATCTCTATGATCGCAATTAATCGAAGTCACTTCAATGTTGAACAACTGGGCAACCTTA 1458  
DB |||||  
QY 1218 CATCTCGTATGATCGGATTAATCGAAGTCACTTCAATGTTGAACAGTTCGGGCAAGCCAA 1277  
DB |||||  
QY 1459 GAAGATGAGCCTGCTGGAGTGTAGCAAGAGGTGTTATTAGAACTGTTACGAAAAAATA 1518  
DB |||||  
QY 1278 GAAGATGAAGTCTCTCGAGTGTGGCAAGAGCGGTTTATCAGAACTGCTCGGAAAAAATA 1337  
DB |||||  
QY 1519 TGGTTGTGTCGAGTGGATTTTGCACAGCCATTTTCTTAAAGGAATATTTAGAAAGCCA 1578  
DB |||||



Db 1338 CGGCTATGTCAGTGAGCTTTGGCAGCCATTTTCTTTGAAGGAATATTAGAAAGGCA 1397  
Qy 1579 AAGTCAGAAACCGGTGCTGCTCTACTTTCCCTCGAGCAAGGGTTGTTACAGACTATACT 1638  
Db 1398 AAGTCAGAAACCTGTATCTGCTCCCTCTCTTTGGAGCAGACACTGTTACCGAATCCT 1457  
Qy 1639 TCCTTCAAGACCCAGTGATGCTGCTGATGAAGGTAGAGACACGTCCTCAATTAATGATCCAG 1698  
Db 1458 TCCTTCAAGACCTGTATGCTGCTGCTGCCGAACATGAAGACATGTCAGTAATGATGTCAG 1517  
Qy 1699 AATGCAACAGATGATCCTTACGAGGAGGTGATTGCAATCTGGCTGAGCATATCT 1758  
Db 1518 AAACCGGAGAGCAAGCTTCGGAAGGAGGCTGATCGCAACCTGGCGGAGCATCTCT 1577  
Qy 1759 ATTCACTGTAGCAAGTCTGTGCCATTTATGTCCACACACATTTGGCTTGCTCTCTCT 1818  
Db 1578 CTTACCGCCAGCAAGTCTCGGCTATCATGTGCCACCACTGTTGGCTGCTGCTCTCT 1637  
Qy 1819 CTACAGACACAGCGAGGAATGATCTCTCCAATTGGTCGAAGACTTCTTTGTGATGAA 1878  
Db 1638 CTACAGACACAGCGAGGAATCCACTCTCCACGCTGGTGGGAAGACTTCTTTGTGATGAA 1697  
Qy 1879 AGAGGAAGTCTGGCTGCTGATTTGACCTGGGTTCTCAGGAATTCAGAAGATGTAGT 1938  
Db 1698 GGAGGAAGTCTGAGTCTCGGATTTTGACCTGGGCTTCTCCGGGAATTCAGAAGATGTAGT 1757  
Qy 1939 AATGATGCTATACAGTCTCGGGAATTTGCTCAATTCACCACTAGCAGGAACGA 1998  
Db 1758 CATGATGCTATTAGCTTCTGGGAATGCTGTCAATTCACCACTAGCAGGAACGA 1817  
Qy 1999 TGAGTTTTTATACCCCCAGACAACTGTCCCATCAGTCTTTCGAACCTCACTCTACAG 2058  
Db 1818 TGAATTTCTTATCTCCAGCACAACTGTCCCGTCCGCTTTTGAACCTCACTCTACAG 1877  
Qy 2059 CAATGGGTACTTCATGCTTTTATCATGAGGCGCATCATAGCTTCGACGCTTTATGCACT 2118  
Db 1878 CAATGGGTACTTCATGCTTTTATCATGAGGCGCATCATAGCTTCGACGCTTTATGCACT 1937  
Qy 2119 TCTGAACAGAGGGGACTGGGGGTCACATGACACCCCACTACCTGATCAGCAGGA 2178  
Db 1938 CAGAAATAGAGGGGTTCCGAGGGTCTGCGAGGGCTTGGCAACCTGATCAGCAGGA 1997  
Qy 2179 GCAGCTGGTGGGAAGGGGCGCAGCTGTGCTTACCTTCTTCCAATGAAGCAGCCTATCTC 2238  
Db 1998 GCAGCTGGTGGGAAGGGCGCAGCTGTGCTTACCTTCTCTATGAGGTACCATTTTC 2057  
Qy 2239 ACTGCTTGGCCAGACATTTTACCAAGTCTGCCATGAAACAGTGAAGAAAGTTTATCCAGTA 2298  
Db 2058 TCTGCCCTGCCAGACATTTTACAGGTTTGTCAAGAGACAGTAGGAAAGTTTATCCAGTA 2117  
Qy 2299 TGGCAATTTTACAGTGGCAGACGATGACCCAGGAAGATATCATGCTTCTGCTGA 2358  
Db 2118 CGGAATTTCTCAGTGGCAGACAGATGACCCAGGAAGTGTGAGTCTGCGCTTGCAGA 2177  
Qy 2359 GCAGCAGTGGGCAAGAGACTTCCAGAACCTTTGTCTTGGAGAAGTGTGAAGAGATGA 2418  
Db 2178 GCAGCAGTGGGAACAGAGACTTCCGAGGCTCTGAACCTGGAGAAGTGTGAAGAGATGA 2237  
Qy 2419 AGACAGTGAATTTGGGGAGGAACAGCGAGATTGCTTACCTGAAGGTGAGCCCAATCCAAGGA 2478  
Db 2238 GCAGCAGTGAATTTGGTAGGAGCAGCTGATGCTTACCTGAAGGTGAGCCAGGCAAGGA 2297  
Qy 2479 GCACACAGATTATCACCTTCTACAGACTCCTTGGGCTTGTGGGAGGCTTACAG 2538  
Db 2298 GCACAGCAATTTCTACAGAGGCTTCTGGGGGCTTCTGGGGGCTTCTGAGAGGCTACAG 2357  
Qy 2539 CTCTGCTGCATTTTGTTCACAACTTTCAGTGGTCTGCTGTTCAGAACCTGAGTATCTGCA 2598  
Db 2358 CTCTGCTGCATTTTGTTCACACCTTTCGGGGCCAGTTCGGGAGTCTGAGTACCTGCA 2417  
Qy 2599 AAGTTGCAAAATACCTAATAACAGAAACAGAAATGTTGAGTATATGCTGAGAG 2658

Db 2418 GAAGCTGCACAGGTACCTTTCTCACAGGACGGAGAGAAACGTCGCGGTGTACGCTGAGAG 2477  
Qy 2659 TGCCACATATTGCTTGTGAAGAATGCTGTGAAATGTTTAAGGATATTGGGGTTTTTCAA 2718  
Db 2478 TGCCACATATTGCTTGTGAAGAATGCTGTGAAATGTTTAAGGACATCGGGGTTTTTCAA 2537  
Qy 2719 GGAGACCAAAACAAAGAGAGTGTCTGTTTATAGAACTGAGCAGCACATTTTCTACCTCAATG 2778  
Db 2538 AGAGACCAAGCAGAAGCGAGCTGTCTTATAGAACTGAGCAGCACATTTTCTACCTCAGGG 2597  
Qy 2779 CAACCGCAAAACCTTCTAGAAATATATCTGAGTTTGTGCTGTCTGAG 2827  
Db 2598 CAGCCGCGAGAGCTCTCTGGAATACATTTCTGAGCTTCGTGCTGTAG 2646

RESULT 6  
US-09-833-381-1961  
; Sequence 1961, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1: Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1961  
; LENGTH: 612  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-833-381-1961

Query Match 19.4%; Score 581.6; DB 10; Length 612;  
Best Local Similarity 99.0%; Pred. No. 1.4e-176;  
Matches 606; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy 1401 TCTCTATGATCGCATTTATCGAAGTCTACTCAATGTTGAACAACCTGGGCAAACTAAGA 1460  
Db 1 TCCGCTATGATCGCATTTATCGAAGTCTACTCAATGTTGAACAACCTGGGCAAACTAAGA 60  
Qy 1461 AGAATGAGAGCTCTGAGAGTGTAGCAAGAGGTGTTATTAGAAATGTTAGCAAACTATG 1520  
Db 61 AGAATGAGAGCTCTGAGAGTGTAGCAAGAGGTGTTATTAGAAATGTTAGCAAACTATG 120  
Qy 1521 GTTGTGTCGAGTGGATTTTGCACAGCCATTTTCTTAAAGGAATATTAGAAAGCCAAA 1580  
Db 121 GTTGTGTCGAGTGGATTTTGCACAGCCATTTTCTTAAAGGAATATTAGAAAGCCAAA 180  
Qy 1581 GTCAAAA - CCGGTGCTGCTCTACTTTCCCTGAGCAAGGTTGTTACAGCTATACTT 1639  
Db 181 GTCAAAAACCGGTGCTGCTCTACTTTCCCTGAGCAAGGTTGTTACAGCTATACTT 240  
Qy 1640 CTTTCAAGACCCAGTGTGCTGTGATGAAGGTAGACACGTCATTAATGAGTCCAGA 1699  
Db 241 CTTTCAAGACCCAGTGTGCTGTGATGAAGGTAGACACGTCATTAATGAGTCCAGA 300  
Qy 1700 AATCAACAGATGAATCCCTACGAAGGAGGTGATTGCAAACTCTGGCTGAGCATATTCTA 1759  
Db 301 AATCAACAGATGAATCCCTACGAAGGAGGTGATTGCAAACTCTGGCTGAGCATATTCTA 360  
Qy 1760 TTCACTGCTAGCAAGTCTGTGCCATTTATGTCACACACATTTGAGTGTGCTGCTCCTC 1819  
Db 361 TTCACTGCTAGCAAGTCTGTGCCATTTATGTCACACACATTTGAGTGTGCTGCTCCTC 420  
Qy 1820 TACAGACACAGGAGGAATGATCTCTCCACATTTGGTCCGAGACTTCTTTGTGATGAAA 1879  
Db 421 TACAGACACAGGAGGAATGATCTCTCCACATTTGGTCCGAGACTTCTTTGTGATGAAA 480  
Qy 1880 GAGGAAGTCTCTGGCTCGTGTGATTTTGTGCTGGGTTCTCAGGAAATTCAGAAAGTGT - AGT 1938



```

; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7822
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-7822

Query Match      12.8%; Score 385.4; DB 11; Length 399;
Best Local Similarity 99.5%; Pred. No. 3.2e-113;
Matches 397; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 630 CTCACACAGACACCGCGGATGGCTTGCAGAGCGCTTCTTACGATCTTTTATTCAAG 689
Db 1 CTCACACAGACACCGCGGATGGCTTGCAGAGCGCTTCTTACGATCTTTTATTCAAG 60

QY 690 AGCGAGATGTGATAGGCGCATGTTTGGCCACCAATGTGACTGAAATGTGC-TGAACAGC 748
Db 61 AGCGAGATGTGATAGGCGCATGTTTGGCCACCAATGTGACTGAAATGTGCTTGNACAGC 120

QY 749 AGTAGATACAGAGCGCAATTGCAGAGTGGCTGTGAATTAAACCCCTGATGGTTCTGCC 808
Db 121 AGTAGATACAGAGCGCAATTGCAGAGTGGCTGTGAATTAAACCCCTGATGGTTCTGCC 180

QY 809 CAGCAGCAATCAAAAGCGCTTAAACAAAGTGAAGAAAGCAAGCTTAAAGATCTTCAAGAA 868
Db 181 CAGCAGCAATCAAAAGCGCTTAAACAAAGTGAAGAAAGCAAGCTTAAAGATCTTCAAGAA 240

QY 869 ATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGACTGGTGGTGGTCTGCTAAAGCTG 928
Db 241 ATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGACTGGTGGTGGTCTGCTAAAGCTG 300

QY 929 TTCAACAGCTCTTTTGGAAACATTCAAATTCACAAAGTCAACTTTGAGATGGTTAAAGCT 988
Db 301 TTCAACAGCTCTTTTGGAAACATTCAAATTCACAAAGTCAACTTTGAGATGGTTAAAGCT 360

QY 989 GCAACTGAGACGAATTTCCCGTCTCTGTTTCTACAGTT 1027
Db 361 GCAACTGAGACGAATTTCCCGTCTCTGTTTCTACAGTT 399

RESULT 10
; Sequence 1964, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1964
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(440)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1964

Query Match      11.6%; Score 348.4; DB 10; Length 440;
Best Local Similarity 98.3%; Pred. No. 3.2e-101;
Matches 349; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

US-09-833-381-1964

; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7822
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-7822

Query Match      12.8%; Score 385.4; DB 11; Length 399;
Best Local Similarity 99.5%; Pred. No. 3.2e-113;
Matches 397; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 630 CTCACACAGACACCGCGGATGGCTTGCAGAGCGCTTCTTACGATCTTTTATTCAAG 689
Db 1 CTCACACAGACACCGCGGATGGCTTGCAGAGCGCTTCTTACGATCTTTTATTCAAG 60

QY 690 AGCGAGATGTGATAGGCGCATGTTTGGCCACCAATGTGACTGAAATGTGC-TGAACAGC 748
Db 61 AGCGAGATGTGATAGGCGCATGTTTGGCCACCAATGTGACTGAAATGTGCTTGNACAGC 120

QY 749 AGTAGATACAGAGCGCAATTGCAGAGTGGCTGTGAATTAAACCCCTGATGGTTCTGCC 808
Db 121 AGTAGATACAGAGCGCAATTGCAGAGTGGCTGTGAATTAAACCCCTGATGGTTCTGCC 180

QY 809 CAGCAGCAATCAAAAGCGCTTAAACAAAGTGAAGAAAGCAAGCTTAAAGATCTTCAAGAA 868
Db 181 CAGCAGCAATCAAAAGCGCTTAAACAAAGTGAAGAAAGCAAGCTTAAAGATCTTCAAGAA 240

QY 869 ATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGACTGGTGGTGGTCTGCTAAAGCTG 928
Db 241 ATGGTTGCCACTGTCTCACCGGCAATGATCAGACTGACTGGTGGTGGTCTGCTAAAGCTG 300

QY 929 TTCAACAGCTCTTTTGGAAACATTCAAATTCACAAAGTCAACTTTGAGATGGTTAAAGCT 988
Db 301 TTCAACAGCTCTTTTGGAAACATTCAAATTCACAAAGTCAACTTTGAGATGGTTAAAGCT 360

QY 989 GCAACTGAGACGAATTTCCCGTCTCTGTTTCTACAGTT 1027
Db 361 GCAACTGAGACGAATTTCCCGTCTCTGTTTCTACAGTT 399

RESULT 11
US-09-833-381-1959
; Sequence 1959, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1959
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1959

Query Match      11.5%; Score 346.6; DB 10; Length 396;
Best Local Similarity 95.9%; Pred. No. 1.1e-100;
Matches 375; Conservative 0; Mismatches 13; Indels 3; Gaps 2;

QY 2178 AGCAGCTGGTGGCGAAGCGCGCCAGCCTGTGCTACCTTCTCTCCAATGAAGGCCACCATCT 2237
Db 1 AGCAGCTGGTGGCGAAGCGCGCCAGCCTGTGCTACCTTCTCTCCAATGAAGGCCACCATCT 60

QY 2238 CACTGCCCTTGCAGACATTTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGT 2297
Db 61 CACTGCCCTTGCAGACATTTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGT 120

QY 2298 ATGGCATTCTTACAGTGGCAGACCATGATCCAGGAGGATATCAGTCTTAGTCTTGGCTG 2357
Db 121 ATGGCATTCTTACAGTGGCAGACCATGATCCAGGAGGATATCAGTCTTAGTCTTGGCTG 180

QY 2358 AGCAGCAGTGGCAACAAGCTTCCAGAACCTTTGTCTTGGAGAGTGTGATGAAGAAGT 2417
Db 181 AGCAGCAGTGGCAACAAGCTTCCAGAACCTTTGTCTTGGAGAGTGTGATGAAGAAGT 240

QY 2418 AAGACAGTACTTTGGGGAGGAAACAGCGAGATGCTACTGAA--GGTAGGCCAATCAA 2475
Db 241 AAGACAGTACTTTGGGGAGGAAACAGCGAGATGCTACTGAAAGGTGAGCCCAATCAA 300
```



Query Match	10.1%;	Score 302;	DB 10;	Length 418;
Best Local Similarity	84.1%;	Pred. No. 3.2e-86;		
Matches 354;	Conservative 0;	Mismatches 61;	Indels 6;	Gaps 1;
Qy	2112	ATGCAGTTCTGAAACAAGAGGGGACTGGGGGGTCCCACTAGACACCCACCTAACTGATCA	2171	
Db	418	ATGCAGTTCTGAAAGAAGGGGNCAGGAGACC-----CGCGTCTCCACAGCTTGGTCA	365	
Qy	2172	GCACGAGCAGCTGGTGCAGGAAGCGGCGCAGCCTGTGCTACTTCTCTCCAAATGAAGGCA	2231	
Db	364	GCCAGGAGCAGCTGGTGCACAAGGCTGCCAGCGTGTCTATCTCTCTCCAATGAAGGCA	305	
Qy	2232	CCATCTCACTGCCTTGCAGACATTTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTA	2291	
Db	304	CCATCTCTCTCCCTGCCAGACCTTTTACCAATTTGCCATGAAAACAGTGGGCGCGTTTA	245	
Qy	2292	TCCAGTATGGCATTTCTTACAGTGGCAGAGCACGATGACCAAGGAGATATCAGTCTCTAGTC	2351	
Db	244	TCCAGTATTTGCAATTTTATAGTTGCCAGCAAGATATACAGGAAGATATCAGTCTCTGTT	185	
Qy	2352	TTGCTGAGCAGCAGTGGGACAAGAAGCTTCCAGAACCTTTGTCTTGAGAAAGTGATGAAG	2411	
Db	184	CTTCCGAGCAGCAGTTGGAACAAGAGTTTCGGNAACCTTTGTCTTGGAAGATGATGAAG	125	
Qy	2412	AAGATGAAGACAGTGAATTTTGGGAGGAAACAGCGAGATTGCTACTCTGAAGGTGAGCCAAT	2471	
Db	124	AAGATGAAGACAGCGATTTTGGTGAGGAGCAACAGAGATTGCTACTCTGAAGGTGAGCCAGT	65	
Qy	2472	CCAAGGAGCACACGACGTTTATCACCTTCTTTACAGACACTCCTTGGGCGCTTCTCTGGAGG	2531	
Db	64	CCAAGGAGCACACGACGTTTATCACCTTCTTGACAGAGGCTCCTCGGCGCTCTCTGCTGAGG	5	
Qy	2532	C 2532		
Db		4 C 4		

```

RESULT 15
US-09-814-353-16182/c
; Sequence 16182, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16182
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-16182

```

Query Match 8.9%; Score 268.6; DB 13; Length 379;  
Best Local Similarity 96.6%; Pred. No. 1.9e-75;

	Matches	285;	Conservative	0;	Mismatches	9;	Indels	1;	Gaps	1;
Qy	1778	TGTGCCATTATGCCACACACATTGTGGCTGCTCTCTACAGACACAGCAGGGA	1837							
Db	379	TGTGCCATTATGCCACACACATTGTGGCTGCTCTCTACAGACACAGCAGGGA	320							
Qy	1838	ATTGATCTCTCCACATTGGTTCGAAGACTCTTTGTGATGAAGAGGAAGTCTGGGCTCGT	1897							
Db	319	AATGATCTCTCCACATTGGTTCGAAGACTCTTTGTGATGAAGAGGAAGTCTGGGCTAGT	260							
Qy	1898	GATTTTGACTCGGGTTCCTCAGGAATTCAGAAAGTGTAGTAAATGCATCCATACAGCTG	1957							
Db	259	AATATTGACCTGGGGTTCCTCAGGAATTCAGAAATGTAGTAAATGCATCCATACAGCTG	200							
Qy	1958	CTGGGAAATGTGTCACAATCACCACTAGCAGGAACGATGAGTTTTTTTATCACCCCC	2017							
Db	199	CTGGGAAATGTGTCACAATCACCC--ACTAGCAGGAATGATGAGTTTTTTATCACCCCC	141							
Qy	2018	AGCACAACTCTCCCATCATGCTTTTGGAACTCAACTTCTACAGCAATGGGTACTTC	2072							
Db	140	AGCACAACTCTCCCATCATGCTTTTGGAACTCAACTTCTACAGCAATGGGTACTTC	86							

Search completed: January 10, 2004, 10:32:02  
Job time : 699.063 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 22:59:07 ; Search time 4035.2 Seconds  
(without alignments)  
18087.435 Million cell updates/sec

Title: US-09-935-290-1  
Perfect score: 3003  
Sequence: 1 ttcggaccaggctgctgcg.....tctgagatcttccagcttt 3003

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:

- 1: em\_estba:\*\*
- 2: em\_esthum:\*\*
- 3: em\_estin:\*\*
- 4: em\_estmu:\*\*
- 5: em\_estov:\*\*
- 6: em\_estpl:\*\*
- 7: em\_estro:\*\*
- 8: em\_htc:\*\*
- 9: gb\_est1:\*\*
- 10: gb\_est2:\*\*
- 11: gb\_est3:\*\*
- 12: gb\_est4:\*\*
- 13: gb\_est5:\*\*
- 14: gb\_est6:\*\*
- 15: em\_estfun:\*\*
- 16: em\_estom:\*\*
- 17: em\_gss\_hum:\*\*
- 18: em\_gss\_inv:\*\*
- 19: em\_gss\_pln:\*\*
- 20: em\_gss\_vrt:\*\*
- 21: em\_gss\_fun:\*\*
- 22: em\_gss\_mam:\*\*
- 23: em\_gss\_mus:\*\*
- 24: em\_gss\_pro:\*\*
- 25: em\_gss\_rod:\*\*
- 26: em\_gss\_phg:\*\*
- 27: em\_gss\_vrl:\*\*
- 28: gb\_gss1:\*\*
- 29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1898.2	63.2	3305	11 AK047295	Mus muscu
2	889.2	29.6	1043	12 BQ062851	AGENCOURT
3	856.6	28.6	1055	12 BQ057957	AGENCOURT
4	848.2	28.2	918	13 BQ437820	AGENCOURT

5	794.8	26.5	968	12 BQ058377	AGENCOURT
6	703.8	23.4	965	10 BF794257	AGENCOURT
7	684	22.8	687	9 AL598680	DKE2p313P
8	582.4	19.4	776	14 CD000106	AGENCOURT
9	573	19.1	573	14 CB269030	H
10	564.2	18.8	637	14 CB268766	H
11	561.4	18.7	753	13 BQ703927	UI-M-FOO-
12	555.6	18.5	799	10 BG704442	602688632
13	547.2	18.2	580	10 AW976326	EST388435
14	530.8	17.7	602	9 AL603562	DKE2p686P
15	530	17.6	988	10 BG034520	602302638
16	529.2	17.6	1137	12 BM471397	AGENCOURT
17	497	16.6	498	13 BQ694826	1000969 H
18	497	16.6	553	9 AL596527	DKE2p451D
19	496.8	16.5	550	13 BQ694822	H
20	490.4	16.3	492	14 CB269165	1008072 H
21	488.6	16.3	587	14 CA406237	1002386 H
22	487.6	16.2	690	12 BI691948	H
23	483.6	16.1	1037	10 BF037399	601460764
C					AA769055 oa74a10.s
24	480	16.0	656	9 AA769055	oa74a10.s
25	467.8	15.6	825	13 BU263739	603815507
26	446	14.9	455	13 BQ694370	1000513 H
27	445	14.8	589	14 CA889049	B0150D03-
28	440.6	14.7	746	14 BY741039	BX741039
29	420.4	14.0	859	12 BI738244	H
C					BM414848 ECC00320
30	398	13.3	446	12 BM414848	603361683
31	395.8	13.2	469	12 BI551617	603194269
32	393	13.1	564	9 AA277375	va81g09.r
33	388.8	12.9	646	12 BI100769	H
34	378.4	12.6	535	14 CA896337	B0198D07-
35	369.6	12.3	527	14 CD554026	B0370D05-
36	368.6	12.3	400	14 R73257	Y109g05.r1
37	366.6	12.2	519	14 CA887606	B0138C12-
38	362.8	12.1	491	14 CA885724	B0122A11-
39	343.4	11.4	347	14 CB265571	H
C					AA172653 ms19c01.r
40	340.2	11.3	571	9 AA172653	ms19c01.r
41	329.8	11.0	411	12 BM146906	TCAA1E75
42	324	10.8	419	12 BM710854	UI-E-DX0-
43	321.8	10.7	823	14 CD362496	AGENCOURT
C					CD502509 CDA57-B03
44	321.4	10.7	1310	14 CD502509	H
45	318	10.6	713	13 BU444448	603766560

ALIGNMENTS

RESULT 1  
AK047295

LOCUS  
DEFINITION

AK047295 3305 bp mRNA linear HTC 05-DEC-2002  
Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length  
enriched library, clone:B930046K04 product:glycerol-3-phosphate  
acyltransferase, mitochondrial, full insert sequence.

ACCESSION  
VERSION

AK047295.1 GI:26092071  
HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

JOURNAL

92279253

MEDLINE

10349636

PUBMED

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE

Normalization and subtraction of cap-trapper-selected cDNAs to

JOURNAL

Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE

20499374

PUBMED

11042159



REFERENCE  
AUTHORS

3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

REFERENCE  
AUTHORS

4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Harada, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,  
Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, F., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,  
Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.  
and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

REFERENCE  
AUTHORS

5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 3305)

REFERENCE  
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akaira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

REFERENCE  
AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/

REFERENCE  
AUTHORS

Location/Qualifiers

## source

1..3305  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:B930046K04"  
/db\_xref="taxon:10090"  
/clone="B930046K04"  
/tissue\_type="cerebellum"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="10 days neonate"  
156..2835  
/note="glycerol-3-phosphate acyltransferase, mitochondrial  
(MGDI|GI:109162, GB|NM\_008149, evidence: BLASTN, 99%,  
match=2883)  
putative"  
3281..3286  
/note="putative"  
3305  
/note="putative"  
BASE COUNT 819 a 861 c 832 g 793 t  
ORIGIN

## Query Match

Best Local Similarity 63.2%; Score 1898.2; DB 11; Length 3305;  
Matches 2229; Conservative 0; Mismatches 445; Indels 15; Gaps 5;

## QY

312 CACATGATTGGGAATTACACATTTGTGACATGGATGAATCTGCACCTGACCCCTTGGTACAA 371

## Db

127 CACATGTTGGGACTTGCACGTTCTGCCATGGAGAGTCTTCAGTGACAGTTGGCACAA 186

## QY

372 TAGATGTTTCTTATCTGCCACATTTTCATCAGAAATACAGATGTTGGTGCATGAAGCACAA 431

## Db

187 TAGAGCTTTCTTATCTGCCACGTTTCATCGGAATACAGCCCTTGGCCGATGTAACACACCA 246

## QY

432 GTGAGAAATGGGGTGAAGTGTGGCTTTAGACCACCGCTTTCAGATCTGCAACTTTAAAT 491

## Db

247 GTGAGGACTGGGTGTGACTGTGGGTTCAACACCTACCTTCTTCAGATCTGCAACACTGAAAT 306

## QY

492 GGAAGAAAGCCTAATGATCGGAAAGGCCAATTTGTTGGAAGATGTTCTTACTCTCTGCA 551

## Db

307 GGAAGAGAGCCTTATGAGCCGGAAGAGGCCCTTCTGTTGGAAGGTGCTCTATTTCTGCA 366

## QY

552 CTCCCAGAGCTGGGACAAATTTTCAACCCAGATATCCCGTCTTTGGGTTTGGCGAATG 611

## Db

367 CGCCACAGAGCTGGGAAGGTTTTCACCCAGATATCCATCTCTGGGTTTGGCGAATG 426

## QY

612 TTATTTATATCAATGAATCTACAGACACCGCGGATGGCTTTCAGACAGCGCTTTCTT 671

## Db

427 TTATTTATATCAATGAATCAACAGGACAGAGGATGGCTGGCGAGACGGCTGTCTT 486

## QY

672 AGTTCTTTTATCAAGAGCGAGATGTGCATTAAGGGCATGTTTGGCCCAATGTGACTG 731

## Db

487 ACATCTCTTTTGTTCAGAGCGAGAGCTCCATTAAGGGCATGTTTGGCCCAATGTGACTG 546

## QY

732 AAAATGTGCTGAACAGCAGTAGAGTACAGAGGCAATTTGCAAGAGTGGCTGCTGAATTA 791

## Db

547 AGATGTACTGAGCAGCAGCAGTCCAGAGGCAATTTGCTGAAGTGGCTGGGAGTTGA 606

## QY

792 ACCGTGATGGTCTGCCCAGAGCAATCAAAAGCGCTTAAAGAGTGAAGAAAGAGCTA 851

## Db

607 ACCAGATGGATCTGCCAGCAGCAGTCCAAAGCCATCCAGAGGTTGAAAGAGAGCA 666

## QY

852 AAGAGTCTTCAAGAAATGTTGCCATGTCTCACCGGCATGTATCAGACTGACTGGGT 911

## Db

667 GGAAGATCTCCAGAGATGGTGGCCACCGTCTCCCAGGATGATCAGCTGACTGGCT 726

## QY

912 GGGTGTCTGTAATAACTGTTTCAACAGCTTCTTTTGGAAACATTCAAATTCAAAGGTCAAC 971

## Db

727 GGGTGTACTAAGCTTTCACAGCTTCTCTGGAACATTCAGATTCAAGAGGTCAAC 786

## QY

972 TTGAGATGTTAAAGCTGCAATGAGACGAATTTGCCGCTTCTGTTTCTACAGTTCA 1031

## Db

787 TCGAGATGTTCAAGGCTGCAACTGAGACGAACCTGCGCTCTTGTCTTCTGCGGTGACA 846



REFERENCE 1 (bases 1 to 1043)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2094 row: 9 column: 16  
High quality sequence stop: 608.  
Location/Qualifiers  
1. 1043  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5924391"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT 296 a 238 c 231 g 278 t

Query Match 29.6%; Score 889.2; DB 12; Length 1043;  
Best Local Similarity 95.7%; Pred. No. 1.1e-241;  
Matches 957; Conservative 0; Mismatches 38; Indels 5; Gaps 4;

QY 212 AGGACAGCGGCTCCCTGTTGATGGACATTTCTGCACCGGAAACTGATAGCTGATGCC 271  
DB 43 AGGACAGCGGCTCCCTGTTGATGGACATTTCTGCACCGGAAACTGATAGCTGATGCC 102

QY 272 TGAAGTTTATGTTATGAACAGAGAACTTTCATCCAGACATGATTTGGGAATTACA 331  
DB 103 TGAAGTTTATGTTATGAACAGAGAACTTTCATCCAGACATGATTTGGGAATTACA 162

QY 332 CTTTGTGACATGATGAATCTGCACCTGACCTGGTACAATAGATGTTCTTATCTGCCA 391  
DB 163 CTTTGTGACATGATGAATCTGCACCTGACCTGGTACAATAGATGTTCTTATCTGCCA 222

QY 392 CATTTCATCAGAAATACAGTGTGTCGATGTAAGACACAAAGTGAAGGATGGGTGATGT 451  
DB 223 CATTTCATCAGAAATACAGTGTGTCGATGTAAGACACAAAGTGAAGGATGGGTGATGT 282

QY 452 GGCTTTAGACCCACCGCTTCAGATCTGCACTTAAATGGAAGAAAGCCTTAATGAGT 511  
DB 283 GGCTTTAGACCCACCGCTTCAGATCTGCACTTAAATGGAAGAAAGCCTTAATGAGT 342

QY 512 CGGAAAGGCGCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAA 571  
DB 343 CGGAAAGGCGCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAA 402

QY 572 TTTTTCACCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAGT 631  
DB 403 TTTTTCACCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAAGT 462

QY 632 CACACAGACACCGCGATGGCTTGAAGCGCTTTCTTACGTTCTTTTATTCAGAG 691  
DB 463 CACACAGACACCGCGATGGCTTGAAGCGCTTTCTTACGTTCTTTTATTCAGAG 522

QY 692 CGAGATGTGCATAAGGCGCATGTTTGCCACCAATGTGACTGAAATGTGCTGAACAGAGT 751

Db 523 CGAGATGTGCATAAGGCGCATGTTTGCCACCAATGTGACTGAAATGTGCTGAACAGAGT 582

QY 752 AGAGTACAGAGGCGCAATTTCAGAGTGGCTGCTGATTAACCCCTGATGTTCTGCCAG 811

Db 583 AGAGTACAGAGGCGCAATTTCAGAGTGGCTGCTGATTAACCCCTGATGTTCTGCCAG 642

QY 812 CAGCAATCAAAAGCGGTTAAACAAAGTGAAGAAAGCTTAAAGGATTTCTTCAAGAAATG 871

Db 643 CAGCAATCAAAAGCGGTTAAACAAAGTGAAGAAAGCTTAAAGGATTTCTTCAAGATATG 702

QY 872 GTTGCCACATGCTCTACCGGCAATGATCAGATGACTGGTGGGTGCTGCTTAAACATGTTTC 931

Db 703 GTTGCCACATGCTCTACCGGCAATGATCAGATGACTGGTGGGTGCTGCTTAAACATGTTTC 762

QY 932 AACAGCTTCTTTTGGGAACATTCAAATTCACAAAGTCAA-CUTTGAAGTGGTTAAAGCTGC 990

Db 763 AACAGCTTCTTTTGGGAACATTCAAATTCACAAAGTCAAACCTTGAGATGGTTAAAGCTGC 822

QY 991 AACTGAGACGAATTTGCCGCTTCTGTTTCTACAGTTTCATAGATCCCATATGACTATCT 1050

Db 823 CACTGAGACGAATTTGCCGCTTCTGTTTCTACAGTTTCATAGATCCCATATGACTATCT 882

QY 1051 -GCTGCTCACTTTCATCTTCTTCCATPAACATCAAGCACCATACATGCTTCAGGCA 1109

Db 883 GGCTGCTCACTTTCATCTTCTTCCGCAATCAATCAAGCACCATACATGCTTCAGGCA 942

QY 1110 ATATCTCAACATCCCAATCTTCAGTACCTTCATCAATGCTT--GGGGGCTTCTTCAT 1167

Db 943 AATATCTCAACATCCCAATCTTCAGTACCTTCATCCCTAAACCTGGGGGGGCTCTCTT 1002

QY 1168 ACACCAAGCGCTCGA-TGAACACACAGATGGACGGAAGA 1206

Db 1003 TCAACAAAGCTTCATTGAAACCCCAATGGACGGTAAAA 1042

RESULT 3  
LOCUS BQ057957  
DEFINITION AGENCOURT\_6924421 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5813732  
5', mRNA sequence.  
ACCESSION BQ057957  
VERSION BQ057957.1 GI:19817297  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1055)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2066 row: d column: 21  
High quality sequence stop: 666.  
Location/Qualifiers  
1. 1055  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5813732"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT	285 a	248 c	240 g	281 t	1 others
ORIGIN					
Query Match	28.6%; Score 858.6; DB 12; Length 1055;				
Best Local Similarity	94.3%; Pred. No. 6.1e-233;				
Matches	945; Conservative	0; Mismatches	50; Indels	7; Gaps	5;
QY	213	GGCAGCAGCGGCTCCCTGTTGTATGGACATCTGACCCGGAACCTGATAGCTGAGTCT	272		
DB	42	GGCAGCAGCGGCTCCCTGTTGTATGGACATCTGACCCGGAACCTGATAGCTGAGTCT	101		
QY	273	GAAGTTTATGTTATGAACAGAGAACTTTCATCCAGCAGATGTTGGGAATTACAC	332		
DB	102	GAAGTTTATGTTATGAACAGAGAACTTTCATCCAGCAGATGTTGGGAATTACAC	161		
QY	333	TTTGTGACATGATGAATCTGACCTGACCTTGGTACATAGATGTTCTTATCTGCCAC	392		
DB	162	TTTGTGACATGATGAATCTGACCTGACCTTGGTACATAGATGTTCTTATCTGCCAC	221		
QY	393	ATTTCATCAGATACAGTGTGTCATGTAAGCACACAGTGGAGAAATGGGTGAGTGTG	452		
DB	222	ATTTCATCAGATACAGTGTGTCATGTAAGCACACAGTGGAGAAATGGGTGAGTGTG	281		
QY	453	GCTTTAGACCCACCGCTTTCAGATCTGCAACTTTAAATGGAAGAAAGCCTTAATGAGTC	512		
DB	282	GCTTTAGACCCACCGCTTTCAGATCTGCAACTTTAAATGGAAGAAAGCCTTAATGAGTC	341		
QY	513	GGAAAGGCCATTGTTGGAAGATGTTTACTCTGCACTCCCGAGAGCTGGGCAAAAT	572		
DB	342	GGAAAGGCCATTGTTGGAAGATGTTTACTCTGCACTCCCGAGAGCTGGGCAAAAT	401		
QY	573	TTTTCAACCCAGTATCCCGTCTTTGGGTTTCGGAATGTTATTTATATCAATGAACCTC	632		
DB	402	TTTTCAACCCAGTATCCCGTCTTTGGGTTTCGGAATGTTATTTATATCAATGAACCTC	461		
QY	633	ACACAAGACACCGCGGATGGCTTGAAGACGCTTTCTTACGTTCTTTTATTCAGAGC	692		
DB	462	ACACAAGACACCGCGGATGGCTTGAAGACGCTTTCTTACGTTCTTTTATTCAGAGC	521		
QY	693	GAGATGTGATGAAGGCATGTTTGCCACCAATGTGACTGMAAATGTGCTGAACAGCTA	752		
DB	522	GAGATGTGATGAAGGCATGTTTGCCACCAATGTGACTGMAAATGTGCTGAACAGCTA	581		
QY	753	GAGTACAAGAGCAATTGACAGAGTGGCTGCTGAATTAACCCCTGATGGTCTGCCACG	812		
DB	582	GAGTACAAGAGCAATTGACAGAGTGGCTGCTGAATTAACCCCTGATGGTCTGCCACG	641		
QY	813	AGCAATCAAAAGCCGTTAAACAAGTGAAGAAAGAAAGCTTCTTCAAGAAATGG	872		
DB	642	AGCAATCAAAAGCCGTTAAACAAGTGAAGAAAGAAAGCTTCTTCAAGAAATGG	701		
QY	873	TTGCCATGTCTCAACCGCAATGATCAGATGACTGGGTGGGTGCTGTAAGAACTGTTCA	932		
DB	702	TTGCCATGTCTCAACCGCAATGATCAGATGACTGGGTGGGTGCTGCTAAACTGTTCA	761		
QY	933	ACAGCTCTTTTGGACATTCAAATTCACAAGGTCACCTGAGATGGTTAAAGCTGCAA	992		
DB	762	ACAGCTCTTTTGGACATTCAAATTCACAAGGTCACCTGAGATGGTTAAAGCTGCAA	820		
QY	993	CTGAGACGAATTTGCGGCTTCTGTTTCTACAGGTTTCATAGATCCCATATGACTATGTC	1052		
DB	821	CTGAGACGAATTTGCGGCTTCTGTTTCTACAGGTTTCATAGATCCCATATGACTATGTC	880		
QY	1053	T-GCTCAGTTTCAATCTCTCTTCTGCAATCAACATCAAGGACCCATACATTTGCTTCAGGCAAT	1111		

Db	881	TGCTCACTTCTCTCTCCGCCATTAACCTAAGGCCCTTTTCATTCGTTTCAGGCAAT	940																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
----	-----	--	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLM2067 row: i column: 02  
 High quality sequence stop: 722.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```

1. - 368
   /organism="Homo sapiens"
   /mol_type="mrna"
   /db_xref="taxon:9606"
   /clone="IMAGE:5614217"
   /tissue_type="lymphoma"
   /lab_host="PH10B (phage-
   /clone_lib="NIH_MGC_99"
   /note="CDNA: lymph; Vec-
   /EcoRI; cDNA made by olig
   into EcoRI/XhoI sites usi
   GGCACGAG(G). Size-select
   1.8kb. Library construct
   of Gerald M. Rubin (Univ
   using ZAP-cDNA synthesiz
   II RT (Life Technologies
   Library."

```

BASE COUNT	276 a	214 c	216 g	262 t
ORIGIN				

Query Match	26.5%	Score 794.8;	DB 12;	Length 968;
Best Local Similarity	97.3%	Pred. No. 9.3e-215;		
Matches 852; Conservative	0;	Mismatches 17;	Indels 7;	Gaps 4;

213	Qy	GGCAGACGGCTCCCTGTTGTATAGCAACATCTGCACCCGAACTCATAGCTGAGTCCCT	272
42	Db	GGCAGCAGCGCTCCCTGTTGTATAGCAACATCTGCACCCGAACTCATAGCTGAGTCCCT	101
273	Qy	GAAGTTTTATGTTATGAAACAGAAGAACTTTTCATCCACGACACATGATTTGGGAATTACAC	332
102	Db	GAAGTTTTATGTTATGAAACAGAAGAACTTTTCATCCACGACACATGATTTGGGAATTACAC	161
333	Qy	TTTGTGATCGATGGATGTAATCTGCACTGACCCCTTGGTACAATAGATGTTTCTTATCTGCCAC	392
162	Db	TTTGTGATCGATGGATGTAATCTGCACTGACCCCTTGGTACAATAGATGTTTCTTATCTGCCAC	221
393	Qy	ATTTCATCAGAAATACAGTGTGGTTCGATGTAAAGCACACAAAGTGAAGGAATGGGGTCAGTGTG	452
222	Db	ATTTCATCAGAAATACAGTGTGGTTCGATGTAAAGCACACAAAGTGAAGGAATGGGGTCAGTGTG	281
453	Qy	GC'TTTAGACCCACCGTCTTCAGATCTGCAACTTTTAAATGGAAAGAAAGCCTAATGAGTC	512
282	Db	GC'TTTAGACCCACCGTCTTCAGATCTGCAACTTTTAAATGGAAAGAAAGCCTAATGAGTC	341
513	Qy	GGAAAAGGCCAATTTGTTGGAAAGATGTTGTTACTCTGCACTCCCCAGAGCTGGGACAAT	572
342	Db	GGAAAAGGCCAATTTGTTGGAAAGATGTTGTTACTCTGCACTCCCCAGAGCTGGGACAAT	401
573	Qy	TTTTTCAACCCAGTATCCCGTCTTTGGGTTTGGGGAATGTTATTATTAATCAATCAAACTC	632
402	Db	TTTTTCAACCCAGTATCCCGTCTTTGGGTTTGGGGAATGTTATTATTAATCAATCAAACTC	461
633	Qy	ACAAAGACACCCGGGATGCGTTTGCAGAGCGCTTTCTTACGTTCTTTTATTAATCAAGAGC	692
462	Db	ACAAAGACACCCGGGATGCGTTTGCAGAGCGCTTTCTTACGTTCTTTTATTAATCAAGAGC	521
693	Qy	GAGATGTGCAATAGGGGATGTTTGGCAACCAATGTGACTGAAAATGTGCTGAAACAGCAGTA	752
522	Db	GAGATGTGCAATAGGGGATGTTTGGCAACCAATGTGACTGAAAATGTGCTGAAACAGCAGTA	581
753	Qy	GAGTACAAGAGGCCAATTCGACAAAGTGGCTGCTGAAATTTAAACCCCTGATGTTTCTGCCACG	812
582	Db	GAGTACAAGAGGCCAATTCGACAAAGTGGCTGCTGAAATTTAAACCCCTGATGTTTCTGCCACG	641
813	Qy	AGCAATCAAAAGCCGTTTAAACAAAGTGAAGAAAGCTTAAAGGATTTCTTCAAGAAATGG	872
642	Db	AGCAATCAAAAGCCGTTTAAACAAAGTGAAGAAAGCTTAAAGGATTTCTTCAAGAAATGG	701

1871	Qy	GTGATGAAGAGGAAGTCTGGCTCGTGATTTTGAACCTGGGGTTCTCAGGAAATTCAGAA	1939
241	Db	GTGATGAAGAGGAAGTCTGGCTCGTGATTTTGAACCTGGGGTTCTCAGGAAATTCAGAA	300
1931	Qy	GATGTAGTAAATGCATGCATACACACTGCTGGGAAATTTGTCACAATCACCCACACTAGC	1990
301	Db	GATGTAGTAAATGCATGCATACACACTGCTGGGAAATTTGTCACAATCACCCACACTAGC	360
1991	Qy	AGGAACGATGAGTTTTTTTATCACCCCCAGACACAACCTGTCCCATCAGTCTTCGAACTCAAC	2050
361	Db	AGGAACGATGAGTTTTTTTATCACCCCCAGACACAACCTGTCCCATCAGTCTTCGAACTCAAC	420
2051	Qy	TTCTACAGCAATGGGGTACTTCATGTCTTTATCATGAGGCCATCATAGCTTGCAGCCTT	2110
421	Db	TTCTACAGCAATGGGGTACTTCATGTCTTTATCATGAGGCCATCATAGCTTGCAGCCTT	480
2111	Qy	TATGCAGTTCTGAACAAGAGGGGACTGGGGGTCCCACTAGCACCCCACTTAACCTGATC	2170
481	Db	TATGCAGTTCTGAACAAGAGGGGACTGGGGGTCCCACTAGCACCCCACTTAACCTGATC	540
2171	Qy	AGCCAGGAGCAGCTGGTGCAGAGCGCCAGCCCTGTGCTACCTTCTCTCCATGAGGC	2230
541	Db	AGCCAGGAGCAGCTGGTGCAGAGCGCCAGCCCTGTGCTACCTTCTCTCCATGAGGC	600
2231	Qy	ACCATCTCACTGCCTTGCAGACATTTTTACCAAGTCTGCCATGAACAGTAGGAAGTTT	2290
601	Db	ACCATCTCACTGCCTTGCAGACATTTTTACCAAGTCTGCCATGAACAGTAGGAAGTTT	660
2291	Qy	ATCCAGTATGSCATTTCTTACAGTGGCAGAGCAGCATCACCAGG-AAAGATATCAGTCCTAG	2349
661	Db	ATCCAGTATGSCATTTCTTACAGTGGCAGAGCAGCATCACCAGGAAAGATATCAGTCCTAG	720
2350	Qy	TCCTTGCTGAGCAGCAGTGGGACAAGAAGCTTCAGAAACCTTT-GTCTTGAGAGAAGTATG	2408
721	Db	TCCTTGCTGAGCAGCAGTGGGACAAGAAGCTTCAGAAACCTTTGCTCTTGAGAGAAGTATG	780
2409	Qy	AAGAAGATGAAGACAGTG-ACCTTTGGGGAGGAACAGCGAGATTGCTACCTGAAGGTGAGC	2467
781	Db	AAGAAGATGAAGACAGTGNACTTTGGGGAGGAACAGCGAGATTGCTACCTGAAGGTGAGC	840
2468	Qy	CAA-TCCAAGGAGCACCAGCAGTTTATCACCTCTTACAGAGACTCCTTGGG--CCTTTG	2524
841	Db	CAANTCCAGGAGCACCAGCAGTTTATCACCTCTTACANAGACTCCTTGGGNCCTTGG	900
2525	Qy	CTGGAGGCCTACAGCTCT	2542
901	Db	CTGGNAGNCTACAGCTCT	918

RESULT 5  
 BQ058377  
 LOCUS  
 DEFINITION  
 BQ058377 968 bp mRNA linear EST 29-MAR-2002  
 AGENCOURT 6794559 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5814217  
 5', mRNA sequence.  
 ACCESSION  
 BQ058377  
 VERSION  
 BQ058377.1 GI:19817717  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 968)  
 AUTHORS  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL  
 Unpublished.  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabps-remail.nih.gov](mailto:cgabps-remail.nih.gov)  
 Tissue Procurement: Lou Staudt  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be

```
QY 873 TTGCCACTGCTCTCACGGCAATGATCAGACTGCTGGGTGGGTGCTGTCTAAACCTGTTC 932
Db 702 TTGCCACTGCTCTCACGGCAATGATCAGACTGCTGGGTGGGTGCTGTCTAAACCTGTTC 761
QY 933 ACAGCTTCTTTTGGAAACATTCAAAATTCACAAAGGTCACCTTGAGATGTTAAAGCTGCAA 992
Db 762 ACAGCTTCTTTTGGAAACATTCAAAATTCACAAAGGTCACCTTGAGATGTTAAAGCTGCAA 821
QY 993 CT--GAGACGAATTTGGCGCTTCTGTTTCTACCAAGTT--CATAGATCCCATATTGACTA 1047
Db 822 CTGGAGACGAATTTGGCGCTTCTGTTTCTACCAAGTTCTATAGATCCCATATTGACTA 881
QY 1048 T-CTCTGCTAC-TTTCATTTCTTCTGCCATAAC 1081
Db 882 TCCTGCTGCTAC-TTTCATTTCTTCTGCCATAAC 917

RESULT 6
BF794257
LOCUS
DEFINITION
602255455F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4338997 5',
mRNA sequence.
ACCESSION
BF794257
VERSION
BF794257.1 GI:12099311
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 965)
AUTHORS
NIH-MGC http://mgc.mci.nih.gov/.
JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LILAM9948 row: m column: 14
High quality sequence stop: 696.
FEATURES
Location/Qualifiers
1..965
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4338997"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT
273 a 222 c 226 g 244 t
ORIGIN
Query Match 23.4%; Score 703.8; DB 10; Length 965;
Best Local Similarity 97.6%; Pred. No. 8.6e-189;
Matches 758; Conservative 0; Mismatches 12; Indels 7; Gaps 4;

QY 1304 GGAAAACTCTTGCTCGGCAGAGACTTTGTGAGTGTGATGACTGTGCTACC 1363
Db 1 GGAAAACTCTTGCTCGGCAGAGACTTTGTGAGTGTGATGACTGTGCTACC 60

QY 1364 AATGTCATCCAGACATCTTGATATACCTGTGGAATCTCTATGATCGCATATTCGAA 1423
Db 61 AATGTCATCCAGACATCTTGATATACCTGTGGAATCTCTATGATCGCATATTCGAA 120
```

```
QY 1424 GGTCACTACATGTGTGAACAACTGGCAAACTTAAGAGAAATGAGAGCTCTGGAGTGA 1483
Db 121 GGTCACTACATGTGTGAACAACTGGCAAACTTAAGAGAAATGAGAGCTCTGGAGTGA 180
QY 1484 GCAAGAGTGTATAGAAATGTTACGAAAAAATCTATGGTGTGTCCGAGTGGATTTGCA 1543
Db 181 GCAAGAGTGTATAGAAATGTTACGAAAAAATCTATGGTGTGTCCGAGTGGATTTGCA 240
QY 1544 CAGCCATTTTCTTTAAAGGAATATTTAGAAAGCCAAAGTTCAGAAACCGGTGTCTGCTCTA 1603
Db 241 CAGCCATTTTCTTTAAAGGAATATTTAGAAAGCCAAAGTTCAGAAACCGGTGTCTGCTCTA 300
QY 1604 CTTTCCCTGGAGCAAGCGTTGTTACCACTATATCTTCTTCAAGACCCAGTGTCTGCT 1663
Db 301 CTTTCCCTGGAGCAAGCGTTGTTACCACTATATCTTCTTCAAGACCCAGTGTCTGCT 360
QY 1664 GATGAAGGTAGAGACACGTCCTCAATTAAGTTCAGAAATGCAACAGATGAATCCCTACGA 1723
Db 361 GATGAAGGTAGAGACACGTCCTCAATTAAGTTCAGAAATGCAACAGATGAATCCCTACGA 420
QY 1724 AGGAGGTGATTGCAAAATCTGGCTGAGCATATTTCTTACTGCTAGCAAGTCTCTGTGCC 1783
Db 421 AGGAGGTGATTGCAAAATCTGGCTGAGCATATTTCTTACTGCTAGCAAGTCTCTGTGCC 480
QY 1784 ATTATGTCACACACATTTGTGGCTTGCTCTCTTACAGACACAGCAGGGAATTGAT 1843
Db 481 ATTATGTCACACACATTTGTGGCTTGCTCTCTTACAGACACAGCAGGGAATTGAT 540
QY 1844 CTCTCCACATTTGTCGAAGACTTCTTTGTGATGAAGAGGAAGTCTCTGGCTCGTGATTTT 1903
Db 541 CTCTCCACATTTGTCGAAGACTTCTTTGTGATGAAGAGGAAGTCTCTGGCTCGTGATTTT 600
QY 1904 GACTGGGGTCTCTCAGGAATTCAGAAAGTGTAGTAATGTCATGCATACACAGTCTGTGGGA 1963
Db 601 GACTGGGGTCTCTCAGGAATTCAGAAAGTGTAGTAATGTCATGCATACAGTCTGTGGGA 660
QY 1964 -AATTGTGTCAATCAATCAACCCACTAGCAGGAACGATGAGTTTATATCAACCCAGCAC 2022
Db 661 CAATTGTGTCAATCAATCAACCCACTAGCAGGAACGATGAGTTTATATCAACCCAGCAC 718
QY 2023 AACTGTCCATCAGTCTTCGAACTCAACTTCTCAGCAATGGGGTACTTTCATGCTTT 2079
Db 719 CACTGTCC--TCAGTCTTCGACTCAAGTTTCTACAGGA--TGGGTACTTTCATGCTTT 771

RESULT 7
AL598680
LOCUS
DEFINITION
DKFZp313P1521_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
DKFZp313P1521 5', mRNA sequence.
ACCESSION
AL598680
VERSION
AL598680.1 GI:15161371
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 687)
AUTHORS
Poustka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann, S.).
TITLE
EST (Poustka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann, S.).
JOURNAL
Unpublished
COMMENT
Contact: Poustka A.J.
Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innestrasse 73, 14195 Berlin, Germany
Tel: +49-30-84131623
Fax: +49-30-84131128
Email: poustka@mpg-berlin-dahlem.mpg.de
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
```

Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.  
No s1 sequence available.  
This clone (DKP2p313P1521) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

Location/Qualifiers  
1. .687  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKP2p313P1521"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="313 (synonym: hlcc2)"  
/notes="vector: pTriplEx2; Site\_1: SfiIA; Site\_2: SfiIB; cDNA-collection"  
202 a 139 c 168 g 178 t

BASE COUNT 202 a 139 c 168 g 178 t  
ORIGIN  
Query Match 22.8%; Score 684; DB 9; Length 687;  
Best Local Similarity 100.0%; Pred. NO. 3.2e-183; Indels 0; Gaps 0;  
Matches 684; Conservative 0; Mismatches 0;

Qy 2220 CCAATGAAGGCACCATCTCACTGCTTGGCCAGACATTTTACCAAGTCTGCCATGAACAG 2279  
Db 1 CCAATGAAGGCACCATCTCACTGCTTGGCCAGACATTTTACCAAGTCTGCCATGAACAG 60  
Qy 2280 TAGAAGATTATCCAGTATGGCATTTTACAGTGGCAGACGACGATGACACAGAGATA 2339  
Db 61 TAGAAGATTATCCAGTATGGCATTTTACAGTGGCAGACGACGATGACACAGAGATA 120  
Qy 2340 TCAGTCTCTGCTGTCAGCAGAGTGGCAGACGATGTCAGACCTTCTCTTGGGA 2399  
Db 121 TCAGTCTCTGCTGTCAGCAGAGTGGCAGACGATGTCAGACCTTCTCTTGGGA 180  
Qy 2400 GAAGTGTATGAAGAGATGAAGACAGTCACTTTTGGGAGGAGAACGACGAGATTGCTACCTGA 2459  
Db 181 GAAGTGTATGAAGAGATGAAGACAGTCACTTTTGGGAGGAGAACGACGAGATTGCTACCTGA 240  
Qy 2460 AGTGAGCCCAATCAAGAGACACAGCAGTGTATTCACCTTCTTACAGAGACTCTCTGGGC 2519  
Db 241 AGTGAGCCCAATCAAGAGACACAGCAGTGTATTCACCTTCTTACAGAGACTCTCTGGGC 300  
Qy 2520 CTTTGTCTGGAGGCTCAGCTCTGCTGCCATCTTGTGTCACACTTCAGTGTCTCTGTC 2579  
Db 301 CTTTGTCTGGAGGCTCAGCTCTGCTGCCATCTTGTGTCACACTTCAGTGTCTCTGTC 360  
Qy 2580 CAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTTAATAACAGACAGAAAGAAATG 2639  
Db 361 CAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTTAATAACAGACAGAAAGAAATG 420  
Qy 2640 TTGCAGTATATGCTGAGAGTGCCACATATTTGCTGTGAAGAAATGCTGTGAAATGTTTA 2699  
Db 421 TTGCAGTATATGCTGAGAGTGCCACATATTTGCTGTGAAGAAATGCTGTGAAATGTTTA 480  
Qy 2700 AGGATATTTGGGTTTTCAAGGAGACCAACAAAGAGAGTGTCTGTTTGAAGACTGACCA 2759  
Db 481 AGGATATTTGGGTTTTCAAGGAGACCAACAAAGAGAGTGTCTGTTTGAAGACTGACCA 540  
Qy 2760 GCACCTTTCTACCTCAATGCAACCGACAAAACCTTCTAGAAATATATTCAGATTTTGTGG 2819  
Db 541 GCACCTTTCTACCTCAATGCAACCGACAAAACCTTCTAGAAATATATTCAGATTTTGTGG 600  
Qy 2820 TGCTGTAGGTAACTGTGGCACTGCTGGCAAAATGAAGGTATGAGATGAGTCTCTGTGTAG 2879  
Db 601 TGCTGTAGGTAACTGTGGCACTGCTGGCAAAATGAAGGTATGAGATGAGTCTCTGTGTAG 660  
Qy 2880 GTACCAGCTTCTGGCTCAAGGTT 2903  
Db 661 GTACCAGCTTCTGGCTCAAGGTT 684

## RESULT 8

CD000106 776 bp mRNA linear EST 01-MAY-2003  
AGENCY: 13642874 NIH\_MGC\_186 Homo sapiens cDNA clone  
IMAGE: 30322827 5', mRNA sequence.  
CD000106  
CD000106.1 GI: 30294625  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 776)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDCM124 row: h column: 04  
High quality sequence stop: 549.  
Location/Qualifiers  
1. .776  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30322827"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: SfiI  
(ggcattatggcc); Site 2: SfiI (ggcgctcgcc); Library is  
oligo-dT primed and directionally cloned. cDNA was  
prepared from a pooled samples of tissues from skin,  
meninges, duramater, pia mater and choroid plexus.  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor  
sequence: 5'-ATTCTAGAGCGCGCGCCGACATG-TT(30)BN-3'  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC  
Library"

## FEATURES

source  
BASE COUNT 208 a 180 c 199 g 189 t  
ORIGIN  
Query Match 19.4%; Score 582.4; DB 14; Length 776;  
Best Local Similarity 95.1%; Pred. NO. 3.3e-154;  
Matches 643; Conservative 0; Mismatches 28; Indels 5; Gaps 4;

Qy 2312 GTGGCAGACGATGACCGAGGATATCAGTCTAGTCTGCTGAGCAGAGTGGAC 2371  
Db 1 GGGGAGAGACGATGACCGAGGATATCAGTCTAGTCTGCTGAGCAGAGTGGAC 60  
Qy 2372 AAGAAGCTTCCAGAACCTTTGTCTTGGAGAGTGTGAAGAAGATGAACAGTGTGACTTT 2431  
Db 61 AAGAAGCTTCCAGAACCTTTGTCTTGGAGAGTGTGAAGAAGATGAACAGTGTGACTTT 120  
Qy 2432 GGGGAGAGACGAGGAGATTTGCTTACCTGAAGTGTGAGCAATCCAGAGACAGAGTGT 2491  
Db 121 GGGGAGAGACGAGGAGATTTGCTTACCTGAAGTGTGAGCAATCCAGAGACAGAGTGT 180  
Qy 2492 ATCACCCTTCTACAGAGACTCTTGGCCCTTTGCTGGAGGCTTACAGTCTCTGTCATC 2551  
Db 181 ATCACCCTTCTACAGAGACTCTTGGCCCTTTGCTGGAGGCTTACAGTCTCTGTCATC 240  
Qy 2552 TTGTGTTCACAACTTTCAGTGTGCTCTGTTCCAGAACCTGTAGTATCTGCAAAAGTTGCAAAA 2611



Db 241 TTTGTTTCAAACTTTCAGTGGTCTCTGTTTCAGAACTGAGTATCTGCAAAAGTTGCACAAA 300  
 Qy 2612 TACCTAATAACACAGAACAAAGAAATGTTGAGTATATGCTGAGAGTGCACATATTGT 2671  
 Db 301 TACCTAATAACACAGAACAAAGAAATGTTGAGTATATGCTGAGAGTGCACATATTGT 360  
 Qy 2672 CTTGTGAAGATGCTGTGAAATGTTTAAAGATATGAGGTTTCAAGGAGACCAACAA 2731  
 Db 361 CTTGTGAAGATGCTGTGAAATGTTTAAAGATATGAGGTTTCAAGGAGACCAACAA 420  
 Qy 2732 AAGAGAGTCTGTTTGAAGTCTGAGCAGCAGTCTTCTACCTCAATGCAACCGACAAAA 2791  
 Db 421 AAGAGAGTCTGTTTGAAGTCTGAGCAGCAGTCTTCTACCTCAATGCAACCGACAAAA 480  
 Qy 2792 CTTCTAGAAATATATCTGAGTCTTGTGCTGAGTAACTGAGTGGCACTGCTGGCAAA 2851  
 Db 481 CTTCTAGAAATATATCTGAGTCTTGTGCTGAGTAACTGAGTGGCACTGCTGGCAAA 540  
 Qy 2852 TGAAGTCTAGATGAGTCTTCTGAGTACCACTTCTGCTCAAGAGTTTGAAGGTG 2911  
 Db 541 TGAAGTCTAGATGAGTCTTCTGAGTACCACTTCTGCTCAAGAGTTTGAAGGTG 600  
 Qy 2912 --CCTTCAGAGGTCAGGCTGCTGCTG--TNCAGAGTGAATCTCTGG--AAGACAGT- 2966  
 Db 601 CCCCTCTCGGGGGGGTGGGGCTGCTGCTTCCCGAAATGATCTCTCGAAAGACAAAGTGC 660  
 Qy 2967 CCTTCTNCTCCATGG 2982  
 Db 661 CCTTCTNCTCCATGG 676

RESULT 9  
 LOCUS CB269030 573 bp mRNA linear EST 20-FEB-2003  
 DEFINITION 1007937 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens  
 ACCESSION CB269030  
 VERSION CB269030.1 GI:28443615  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 AUTHORS Yang, R.-Z., Shuldiner, A. and Gong, D.-W.  
 TITLE EST analysis of human adipose gene expression  
 JOURNAL Unpublished  
 COMMENT Contact: Gong Da-Wei  
 Division of Endocrinology, Diabetes and Nutrition  
 University of Maryland  
 660 Redwood St, HH497, Baltimore, MD 21201, USA  
 Tel: 410 706 1672  
 Fax: 410 706 1622  
 Email: dgong@medicine.umaryland.edu  
 PCR Primers  
 FORWARD: CTCGGGAAGCGCCCATTTGTTGTTGTT  
 BACKWARD: AATACGACTACTATAGGCGGAATTGG  
 Seq primer: GTTGTACCCGGGAATTC.

FEATURES  
 source  
 1..573  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /sex="Male and Female"  
 /tissue\_type="Adipose"  
 /clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"  
 /note="Vector: lambdaTriplex"

BASE COUNT 160 a 127 c 124 g 162 t  
 ORIGIN

Query Match 19.1%; Score 573; DB 14; Length 573;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-151;

Matches 573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 133 GTCTTCTCCAGAGTCTCCATCAGCTTTGCTAATCGACTGATTGGAATAAATTTCTCTCAAC 192  
 Db 1 GTCTTCTCCAGAGTCTCCATCAGCTTTGCTAATCGACTGATTGGAATAAATTTCTCTCAAC 60  
 Qy 193 ACCACCAAGTCAAGGATACAGGACAGCGGCTCCCTGTTTGTATGGAACACTTTTCATCCAC 252  
 Db 61 ACCACCAAGTCAAGGATACAGGACAGCGGCTCCCTGTTTGTATGGAACACTTTTCATCCAC 120  
 Qy 253 GAAACTGATAGTCTGAGTCTGAACTTTTATGTTATGAAACAGAAAGAACTTTTCATCCAC 312  
 Db 121 GAAACTGATAGTCTGAGTCTGAACTTTTATGTTATGAAACAGAAAGAACTTTTCATCCAC 180  
 Qy 313 ACATGATTTGGGAATTTACACTTTTGTGACATGGAATGTAATCTGCATGACCTTTGGTACA 372  
 Db 181 ACATGATTTGGGAATTTACACTTTTGTGACATGGAATGTAATCTGCATGACCTTTGGTACA 240  
 Qy 373 AGATGTTTCTTATCTGCCACATTTCAATCAGATACAGTGTGTTGTCGATGTAAGCACACA 432  
 Db 241 AGATGTTTCTTATCTGCCACATTTCAATCAGATACAGTGTGTTGTCGATGTAAGCACACA 300  
 Qy 433 TGAGGAATGGGTGAGTGTGGCTTTAGACCCACCGTCTTTCAGATCTGCAACTTTAAATG 492  
 Db 301 TGAGGAATGGGTGAGTGTGGCTTTAGACCCACCGTCTTTCAGATCTGCAACTTTAAATG 360  
 Qy 493 GAAAGAAAGCCTAATGAGTGGAAAGGCCATTTGTTGGAAGATGTTGTTACTCTCTGCAC 552  
 Db 361 GAAAGAAAGCCTAATGAGTGGAAAGGCCATTTGTTGGAAGATGTTGTTACTCTCTGCAC 420  
 Qy 553 TCCCCAGAGCTGGGACAAAATTTTCAACCCAGATATCCCGTCTTTGGGTTTGCAGAAATG 612  
 Db 421 TCCCCAGAGCTGGGACAAAATTTTCAACCCAGATATCCCGTCTTTGGGTTTGCAGAAATG 480  
 Qy 613 TATTATATCATGAATACACAGACACCGCGATGGCTTTCAGAGACGCGCTTTCTTA 672  
 Db 481 TATTATATCATGAATACACAGACACCGCGATGGCTTTCAGAGACGCGCTTTCTTA 540  
 Qy 673 CGTCTTTTATTCAAGAGCGAGATGTGCATAA 705  
 Db 541 CGTCTTTTATTCAAGAGCGAGATGTGCATAA 573

RESULT 10  
 LOCUS CB268766  
 DEFINITION 1007672 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens  
 ACCESSION CB268766  
 VERSION CB268766  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 AUTHORS Yang, R.-Z., Shuldiner, A. and Gong, D.-W.  
 TITLE EST analysis of human adipose gene expression  
 JOURNAL Unpublished  
 COMMENT Contact: Gong Da-Wei  
 Division of Endocrinology, Diabetes and Nutrition  
 University of Maryland  
 660 Redwood St, HH497, Baltimore, MD 21201, USA  
 Tel: 410 706 1672  
 Fax: 410 706 1622  
 Email: dgong@medicine.umaryland.edu  
 PCR Primers  
 FORWARD: CTCGGGAAGCGCCCATTTGTTGTTGTT  
 BACKWARD: AATACGACTACTATAGGCGGAATTGG  
 Seq primer: GTTGTACCCGGGAATTC.

FEATURES  
 source  
 1..637  
 /organism="Homo sapiens"

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="Male and female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/note="Vector: lambdaTriPEx"
BASE COUNT    173 a   145 c   131 g   184 t       4 others
ORIGIN
Query Match    18.8%; Score 564.2; DB 14; Length 637;
Best Local Similarity 98.4%; Pred. No. 4.7e-149;
Matches 599; Conservative 0; Mismatches 7; Indels 3; Gaps 3;
QY 1002 ATTTGCCGCTCTGTTCTACCGTTCATAGATCCCATATTTGCTATCTGCTGCTCACTT 1061
Db 1 ATTTGCCGCTCTGTTCTACCGTTCATAGATCCCATATTTGCTATCTGCTGCTCACTT 60
QY 1062 TCATTCTCTCTGCTGCAATCAACAGCACCATACATTTGCTTCAGGCAATAATCTCAACA 1121
Db 61 TCATTCTCTCTGCTGCAATCAACAGCACCATACATTTGCTTCAGGCAATAATCTCAACA 120
QY 1122 TCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTCATACGACGAAGGCTCG 1181
Db 121 TCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTCATACGACGAAGGCTCG 180
QY 1182 ATGAACACCATGAGGAGGAGAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATAG 1241
Db 181 ATGAACACCATGAGGAGGAGAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATAG 240
QY 1242 TTGAATTTACTTCGACAGCAGCAATTTCTTGGAGATCTTCTGGAAGGACACAGTCTTAGGA 1301
Db 241 TTGAATTTACTTCGACAGCAGCAATTTCTTGGAGATCTTCTGGAAGGACACAGTCTTAGGA 300
QY 1302 GTGGAATAAACCCTTTGTCTGGGACGAGCTTTTGTTCAGTGTGGTAGACTCTGTCTA 1361
Db 301 GTGGAATAAACCCTTTGTCTGGGACGAGCTTTTGTTCAGTGTGGTAGACTCTGTCTA 360
QY 1362 CCAATGTCATCCAGACATCTTGATATACCTGTGGAATCTCTATGATCGCATATCG 1421
Db 361 CCAATGTCATCCAGACATCTTGATATACCTGTGGAATCTCTATGATCGCATATCG 420
QY 1422 AAGTCTACTCAATGGTGAACAACTGGGCAAACTTAAGAAGATAGAGGCTGTGGAGTG 1481
Db 421 AAGNCACTACATGNGAACAACTGGGCCAACTTAANAGATGAGAACTGTGGAGTG 480
QY 1482 TAGCAAGAGGTGTATT-AGAATGTTACGAAAAAATATGGTGTGTGTCGAGTGGATTTT 1540
Db 481 TANCAGAGGTGTATTAAAGATGTTACGAAAAAATATGGTGTGTGTCGAGTGGATTTT 540
QY 1541 GCACAGCATTTTCTTAAAGGATATTAGAAAGCAAAAGTTCAGAAACCGGTGTCTGCT 1600
Db 541 GCACAGCATTTTCTTAAAGGATATTAGAAAGCC-AAAGTCAGAAACCGG-GGTGCT 598
QY 1601 CTACTTTCC 1609
Db 599 CTACTTTCC 607

RESULT 11
BU703927
LOCUS
DEFINITION BU703927 753 bp mRNA linear EST 09-OCT-2002
IMAGE: 6406167 5', mRNA sequence.
ACCESSION BU703927
VERSION BU703927.1 GI:23631449
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 753)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```

Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs@mail.nih.gov  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

Seq primer: pYX-5.  
Location/Qualifiers  
1. .753  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE: 6406167"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 12.5dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH BMAP F00"  
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaudo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is TAGAGAGCC. This library was created for the University  
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
program coordinator." 196 a 202 c 173 g 180 t 2 others

BASE COUNT 196 a 202 c 173 g 180 t 2 others  
ORIGIN

Query Match 18.7%; Score 561.4; DB 13; Length 753;  
Best Local Similarity 84.9%; Pred. No. 3.2e-148;  
Matches 639; Conservative 0; Mismatches 113; Indels 1; Gaps 1;

QY 1373 CCAGACATCTTGATATACCTGTGGAAATCTCTATGATCGCATATCGAAGTCACTAC 1432  
Db 1 CCGACATCTCTGTCATACCCGCTGGGCATCTCGTATGATCGCATATCGAAGTCACTAC 60  
QY 1433 AATGTTGAACAA-CTGGGCAAACTTAAGAAGATAGAGCCTGTGGAGTGTAGCAAGG 1491  
Db 61 AATGGCAACAAAGTTGGGAAAGCCCAAGAAGAACCGAGAGCCTCTGGAGTGTGGCGAGG 120  
QY 1492 TGTATTAGAAATGTTTACGAAAAAACTATGTTGTGTCGAGTGGATTTTGGCAGGCAAT 1551  
Db 121 CGTTATCAGATGCTCGGAAAAAACTACCGCTACGTCGAGTGGATTTTGGCAGGCAAT 180  
QY 1552 TTCTTTAAAGGAATATTTAGAAAGCCAAAGTCCAGAAACCGGTGTCTGCTCTACTTTCCCT 1611  
Db 181 TTCTTTAAAGGAATATTTAGAAAGCCAGAGTCCAGAAACCTGTATCTGCCCCCTTTCTCT 240  
QY 1612 GGAGCAAGGTTGTTACGAGCTATCTTCTTCAAGACCCAGTGTGCTGCTGATCAAGG 1671  
Db 241 GGAGCAAGCACTGTTTACCGCATCTCTCTTCAAGACCCGAATGATGTTGCTGATGAACA 300  
QY 1672 TAGAGACAGCTCCATTATTAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGGAGTT 1731  
Db 301 TCAGACCTATTCAGTACAGTCCAGAAACCCAGCAGACGAGCCTTCGGACGAGGCT 360  
QY 1732 GATTGCAAACTGGCTGAGCATATTTCTATTCTACTGCTAGCAAGTCTGTGCCATTAATGTC 1791  
Db 361 GATTGCAAACTGGCTGAGCATATTTCTTCTTCAACCGCAGCAAGTCTCGGCTATCATGTC 420



## COMMENT

Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johnq@tigr.org  
Plate: 359

Seq primer: Reverse.

## FEATURES

Location/Qualifiers  
1..580  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="MAGE resequences, MAGN"  
/note="Vector: pBluescriptSKm"

BASE COUNT 174 a 119 c 138 g 149 t

## ORIGIN

Query Match 18.2%; Score 547.2; DB 10; Length 580;

Best Local Similarity 97.7%; Pred. No. 3.2e-144;

Matches 555; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 2229 GCACCATCTCAGTCCTGCGCAGACATTTTACCAAGTCTGCCATGAACAGTAGGAAAGT 2288

Db 1 GCTCCATCTCAGTCCTGCGCAGACATTTTACCAAGTCTGCCATGAACAGTAGGAAAGT 60

Qy 2289 TTATCCAGTAGGCAATTTTACAGTGGCAGACGACGATGACAGGAGATATCAGTCCTA 2348

Db 61 TTATCCAGTAGGCAATTTTACAGTGGCAGACGACGATGACAGGAGATATCAGTCCTA 120

Qy 2349 GTCTTGCTGAGCAGCAGTGGGACAAAGAGCTTCCAGAACCTTTGTCTTGGAGAGTGATG 2408

Db 121 GTCTTGCTGAGCAGCAGTGGGACAAAGAGCTTCCAGAACCTTTGTCTTGGAGAGTGATG 180

Qy 2409 AAGAAGTAGAAGCAGTACTTTGGGAGGAGAACAGCAGAGATTGCTACCTGAAGTGAGCC 2468

Db 181 AAGAAGTAGAAGCAGTACTTTGGGAGGAGAACAGCAGAGATTGCTACCTGAAGTGAGCC 240

Qy 2469 AATCCAGGAGCAGCAGCAGTATACCTTTTACAGAGACTCTTGGGCCCTTTGCTGG 2528

Db 241 AATCCAGGAGCAGCAGCAGTATACCTTTTACAGAGACTCTTGGGCCCTTTGCTGG 300

Qy 2529 AGGCCTACAGCTCTGCTGCCATCTTTGTTCAACACTTTCAGTGGTCTGTTCCAGAACCTG 2588

Db 301 AGGCCTACAGCTCTGCTGCCATCTTTGTTCAACACTTTCAGTGGTCTGTTCCAGAACCTG 360

Qy 2589 AGTATCTGAAAAGTTGCCAATACTTAATACAGGAGAACAGAAATGTTGCGAGTAT 2648

Db 361 AGTATCTGAAAAGTTGCCAATACTTAATACAGGAGAACAGAAATGTTGCGAGTAT 420

Qy 2649 ATGCTGAGTGGCCACATATTTGTTGTAAGAAATGCTGTGAAAATGTTTAAAGGATATG 2708

Db 421 ATGCTGAGTGGCCACATATTTGTTGTAAGAAATGCTGTGAAAATGTTTAAAGGATATG 480

Qy 2709 GGGTTTTTCAAGGAGACCAAAACAAAGAGAGTGTCTGTTTATAGAACTGAGCAGCACTTTTC 2768

Db 481 GGGTTTTTCAAGGAGACCAAAACAAAGAGAGAGGGCTGTTTATAGAACTGAGCAGCACTTTTC 540

Qy 2769 TACTCAATGCAACCGACAAAACTTCT 2796

Db 541 TACTCAATGCAACCGACAAAAATTTT 568

## RESULT 14

AL603562

LOCUS

DEFINITION DKFZp686P027 r1 686 (synonym: hlcc3) Homo sapiens cDNA clone

AL603562 602 bp mRNA linear EST 14-AUG-2001

AL603562 DKFZp686P027 5', mRNA sequence.

AL603562.1 GI:15167068

EST.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 602)  
Ansorge W., Wirkner U., Mewes W., Weil B. and Wiemann S.  
EST (Ansorge, W., Wirkner, U., Mewes, H.W., Weil, B. and Wiemann, S.)  
Unpublished  
Contact: Ansorge W  
MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by EMBL (European Molecular Biology Laboratories,

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.

No si sequence available.

This clone (DKFZp686P027) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES  
source

Location/Qualifiers

1..602

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="DKFZp686P027"

/tissue\_type="human skeletal muscle"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="686 (synonym: hlcc3)"

/note="Vector: pTriplex2; Site\_1: SfiI; Site\_2: SfiIB;

cDNA-collection"

BASE COUNT 159 a 130 c 150 g 163 t

## ORIGIN

Query Match 17.7%; Score 530.8; DB 9; Length 602;

Best Local Similarity 99.6%; Pred. No. 1.5e-139;

Matches 532; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 213 GGCAGCAGCGGCTCCCTGTGTATGGACATCTGCACCGGAACTGTAGCTGAGTCCT 272

Db 69 GGCACAGCGGCTCCCTGTGTATGGACATCTGCACCGGAACTGTAGCTGAGTCCT 128

Qy 273 GAAGTTTTATGTTATGAAACAGAGAACTTTTCATCCAGCACATGATTTGGGAATTACAC 332

Db 129 GAAGTTTTATGTTATGAAACAGAGAACTTTTCATCCAGCACATGATTTGGGAATTACAC 188

Qy 333 TTTGTGACATGATGATCTGCACCTCTGGTACATAGATGTTTCTTATCTGCCAC 392

Db 189 TTTGTGACATGATGATCTGCACCTCTGGTACATAGATGTTTCTTATCTGCCAC 248

Qy 393 ATTTCATCAGATACAGTGTGTCGATGTAAGCACACAAGTGAAGGAGTGGGTGAGTGTG 452

Db 249 ATTTCATCAGATACAGTGTGTCGATGTAAGCACACAAGTGAAGGAGTGGGTGAGTGTG 308

Qy 453 GCTTTAGACCCACCGCTTTCAGATCTGCAACTTTTAAATGGAAAGAAAGCTTAATGAGTC 512

Db 309 GCTTTAGACCCACCGCTTTCAGATCTGCAACTTTTAAATGGAAAGAAAGCTTAATGAGTC 368

Qy 513 GGAAGAGGCAATTTGTTGGAGAGATGTTGATCTCTGCACTCCCGAGAGCTGGGACAAAT 572

Db 369 GGAAGAGGCAATTTGTTGGAGAGATGTTGATCTCTGCACTCCCGAGAGCTGGGACAAAT 428

Qy 573 TTTTCAACCCCGATATCCCGTCTTTGGGTTTGGGAAATGTTATTTATATCAATGAAGTCT 632

Db 429 TTTTCAACCCCGATATCCCGTCTTTGGGTTTGGGAAATGTTATTTATATCAATGAAGTCT 488

Qy 633 ACACAGACACACCGGATGCTTTCGAAGACGCTTCTTACGTTCTTTTATTTCAAGAGC 692

Db 489 ACACAGACACACCGGATGCTTTCGAAGACGCTTCTTACGTTCTTTTATTTCAAGAGC 548

Qy 693 GAGATGTGATTAAGGGCATGTTTGGCCCAATGTGACTGAAATGCTGGAACA 746

Db 549 GAGATGTGATTAAGGGCATGTTTGGCCCAATGTGACTGAAATGCTGGAACA 602

RESULT 15  
BG034520  
LOCUS  
DEFINITION BG034520 988 bp mRNA linear EST 24-JAN-2001  
602302638F1 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:4404225 5',  
mRNA sequence.  
ACCESSION BG034520  
VERSION BG034520.1 GI:12427920  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 988)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10115 row: k column: 10  
High quality sequence stop: 604.  
Location/Qualifiers  
1..988  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4404225"  
/tissue\_type="mammary adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_87"  
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.383 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

FEATURES  
source

BASE COUNT 269 a 250 c 248 g 221 t  
ORIGIN  
Query Match 17.6%; Score 530; DB 10; Length 988;  
Best Local Similarity 93.8%; Pred. No. 3.3e-139;  
Matches 619; Conservative 0; Mismatches 30; Indels 11; Gaps 6;  
Qy 1970 GTCACAATCACCCACACTAGCAGGAACGATGAGTTTATACCCCCAGCACAACTGTC 2029  
Db 1 GTCACAATCACCCACACTAGCAGGAATGATGAGTTTATACCCCCAGCACAACTGTC 60  
Qy 2030 CCATCAGTCTTCGAACCTTCTACAGCAATGGGGTACTGCTTTATCATGGAG 2089  
Db 61 CCATCAGTCTTCGAACCTTCTACAGCAATGGGGTACTGCTTTATCATGGAG 120  
Qy 2090 GCCATCATGCTGCGAGCTTTATGAGTCTGAAACAGAGGGGACTGGGGGTCCCACT 2149  
Db 121 GCCATCATGCTGCGAGCTTTATGAGTCTGAAACAGAGGGGACTGGGGGTCCCACT 180  
Qy 2150 AGCACCCCACTTAACCTGATCAGCCAGGAGCAGTGGTGGCAAGGGCGCCAGCTGTGC 2209  
Db 181 AGCACCCCACTTAACCTGATCAGCCAGGAGCAGTGGTGGCAAGGGCGCCAGCTGTGC 240  
Qy 2210 TACCTTCTCTCAATGAAGGCACCAT-CTCACTGCTTTGCCAGACATTTTACCAGTCTG 2268  
Db 241 TACCTTCTCTCAATGAAGGCACCATCTCACTGCTTTGCCAGACATTTTACCAGTCTG 300  
Qy 2269 CCATGAACAGTAGGAAAGTTTATCCAGTATGGGATCTTACAGTGGCAGACGATGA 2328  
Db 301 CCATGAACAGTAGGAAAGTTTATCCAGTATGGGATCTTACAGTGGCAGACGATGA 360

Qy 2329 CCAGGAAGATATCAGTCTCTAGTCTTGTGAGCAGTGGCAAGAAGCTTCCAGAACC 2388  
Db 361 CCAGGAAGATATCAGTCTCTAGTCTTGTGAGCAGTGGCAAGAAGCTTCCAGAACC 420  
Qy 2389 TTTGTCTTTGGAGAAAGTGAAGAAGATGAAGACAGTGACTTTGGGGAGGAAACAGCGAGA 2448  
Db 421 TTTGTCTTTGGAGAAAGTGAAGAAGATGAAGACAGTGACTTTGGGGAGGAAACAGCGAGA 480  
Qy 2449 TTGCTACTGAAGTGAAGCCCAAT-CCAAGGAGCACCAG-CAGTTTATCACCTTCTTACAG 2506  
Db 481 TTGCTACTGAAGTGAAGCCCAATCCCAAGGAGCACCAGCCAGCTTTATCAACCTTCTTAC 540  
Qy 2507 AGA----CTCCTTGGGCTTTGTGAGGCGCTACA--GCTCTGCTGCCATCTTTGTTTAC 2560  
Db 541 AGAGGACTCCTTTGGGCTTTGTGAGGCGCTACAAGCTCTTGTGCCATCTTTGTTTAC 600  
Qy 2561 AACTT--CAGTGGTCTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTAA 2618  
Db 601 AACTTTCAGTGGTTCTCTGTTTCCAGAACCTTGAATCTGCAAAAGTTGCAAAATACCTTA 660

Search completed: January 10, 2004, 06:10:19  
Job time : 4042.2 secs

**THIS PAGE BLANK (UNCLASSIFIED)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2004, 16:35:17 ; Search time 25 Seconds  
(without alignments)  
1557.524 Million cell updates/sec

Title: US-09-935-290-2

Perfect score: 4280

Sequence: 1 MDESALTGLTGVSLPHSS.....FLPQCNKQLLEYILSFVVL 828

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4270	99.8	828	1 PLSB_HUMAN	Q9hcl2 homo sapien
2	3985.5	93.1	827	1 PLSB_MOUSE	Q61586 mus musculu
3	3955	92.4	828	1 PLSB_RAT	P97564 rattus norv
4	678	15.8	718	1 PLSB_CABEL	Q22949 caenorhabdi
5	505	11.8	870	1 PLSB_XYLEFA	Q9pej7 xylella fas
6	462	10.8	834	1 PLSB_PSEAE	Q9hvw7 pseudomonas
7	452	10.6	886	1 PLSB_XANCP	Q8p3e3 xanthomonas
8	449	10.5	885	1 PLSB_XANAC	Q8pes0 xanthomonas
9	433	10.1	678	1 DAPT_RAT	P98192 mus musculu
10	424	9.9	678	1 DAPT_MOUSE	P98171 rattus norv
11	399.5	9.3	809	1 PLSB_VIBVU	Q8dd48 vibrio vuln
12	398.5	9.3	808	1 PLSB_VIBPA	Q87kn0 vibrio para
13	397	9.3	806	1 PLSB_SALTU	Q8zt16 salmonella
14	397	9.3	806	1 PLSB_SALTU	Q8zkx9 salmonella
15	396	9.3	811	1 PLSB_VIBCH	Q9kvp8 vibrio chol
16	395	9.2	810	1 PLSB_HAEIN	P44857 haemophilus
17	394	9.2	680	1 DAPT_HUMAN	O15228 homo sapien
18	391.5	9.1	806	1 PLSB_ECOLI	P00482 escherichia
19	391	9.1	825	1 PLSB_YERPE	Q8zj18 yersinia pe
20	390.5	9.1	806	1 PLSB_ECO57	P58130 escherichia
21	377	8.8	809	1 PLSB_PASMU	Q9cln7 pasteurella
22	289	6.8	775	1 PLSB_MYCLE	O9x7b0 mycobacteri
23	245	5.7	789	1 PLSB_MYCTU	O53207 mycobacteri
24	229	5.4	621	1 PLS1_MYCTU	Q10775 mycobacteri
25	140	3.3	3144	1 HD_HUMAN	P42858 homo sapien
26	128.5	3.0	967	1 CGE1_SCHPO	O13671 schizosacch
27	124.5	2.9	3110	1 HD_RAT	P51111 rattus norv
28	123.5	2.9	1010	1 UFD2_SCHPO	Q9he05 schizosacch
29	119.5	2.8	1174	1 YJ11_YEAST	P47108 saccharomyc
30	118.5	2.8	1176	1 RPB2_DROME	P08266 drosophila
31	118	2.8	1258	1 SAI_HUMAN	Q8wvm7 homo sapien
32	118	2.8	1258	1 SAI_MOUSE	Q9d3e6 mus musculu
33	116.5	2.7	958	1 BP28_MACFA	Q9gm44 macaca fascu

RESULT 1

ID	PLSB_HUMAN	STANDARD;	PRT;	828 AA.
AC	Q9HCL2;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glycerol-3-phosphate acyltransferase, mitochondrial precursor			
DE	(EC 2.3.1.15) (GPAT).			
GN	GPAT OR KIAA1560.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Ashwell R.;			
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE OF 167-828 FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=20450683; PubMed=10997877;			
RA	Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes.			
RT	XVIII. The complete sequences of 100 new cDNA clones from brain which			
RL	code for large proteins in vitro.";			
CC	DNA Res. 7:273-281(2000).			
CC	-!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-			
CC	acyl-en-glycerol 3-phosphate.			
CC	-!- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS. IT MAY			
CC	ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial			
CC	outer membrane (By similarity).			
CC	-!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
CC	EMBL; AL391986; -; NOT ANNOTATED_CDS.			
DR	EMBL; AB046780; BAB13356.1; -;			
DR	MIM; 602395; -;			
DR	InterPro; IPR002123; Acyltransferase.			
DR	Pfam; PF01553; Acyltransferase; 1.			
DR	SMART; SM00563; P1sc; 1.			
KW	Phospholipid biosynthesis; Transferase; Acyltransferase;			
KW	Transmembrane; Mitochondrion; Transit peptide.			
FT	TRANSIT 1 ? MITOCHONDRION (POTENTIAL).			
FT	CHAIN ? 828 GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.			
FT	DOMAIN ? 471 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).			
FT	TRANSMEM 472 494 POTENTIAL.			
FT	DOMAIN 495 574 CYTOPLASMIC (POTENTIAL).			

34	115.5	2.7	3119	1	HD_MOUSE	P42859 mus musculu
35	115	2.7	3744	1	YHP9_YEAST	P38811 saccharomyc
36	111	2.6	987	1	K6P1_YEAST	P16861 saccharomyc
37	111	2.6	1162	1	SA2_HUMAN	Q8n3u4 homo sapien
38	110.5	2.6	820	1	MUTS_CHLTR	O84797 chlamydia t
39	110.5	2.6	1024	1	GCPS_HUMAN	Q96rt8 homo sapien
40	109	2.5	1235	1	KPB2_RABIT	P46018 oryctolagus
41	108.5	2.5	768	1	LIPS_RAT	P15304 rattus norv
42	108	2.5	520	1	GUAA_RHILO	Q987r3 rhizobium l
43	108	2.5	814	1	AKH_BUCAP	Q8k9u9 buchnera ap
44	108	2.5	1162	1	SA2_MOUSE	O35638 mus musculu
45	108	2.5	2273	1	HPA1_YEAST	P32874 saccharomyc

ALIGNMENTS



```
FT TRANSMEM 575 593 POTENTIAL.
FT DOMAIN 594 828 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
SQ SEQUENCE 828 AA; 93836 MW; 55CD321A23D0B65B CRC64;

Query Match
Best Local Similarity 99.8%; Score 4270; DB 1; Length 828;
Matches 826; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDESALTGLTIDVSYLPHSSEYSGRCCKHTSEWEGCGFRPTVFRSATLKWESLSMSRKR 60
DB 1 MDESALTGLTIDVSYLPHSSEYSGRCCKHTSEWEGCGFRPTVFRSATLKWESLSMSRKR 60
QY 61 PFVGRCCYCTPQSDWKFFNPSISGLRNVIYNETHRHWGLARRLSYVLFQERDV 120
DB 61 PFVGRCCYCTPQSDWKFFNPSISGLRNVIYNETHRHWGLARRLSYVLFQERDV 120
QY 121 HKGMFATNTENVNLSRVQEAIAEVAALNPDGSAQQSKAVNKKAKRILOEMVAT 180
DB 121 HKGMFATNTENVNLSRVQEAIAEVAALNPDGSAQQSKAVNKKAKRILOEMVAT 180
QY 181 VSPAMIRLTGWLLKLFNSFFWNIQHKGLEVMVKAATETNPLFLPVHRSHIDYLLT 240
DB 181 VSPAMIRLTGWLLKLFNSFFWNIQHKGLEVMVKAATETNPLFLPVHRSHIDYLLT 240
QY 241 FILFCHNIKAPIYASGNLNIPFSTLIHKLGGFFIRRRIDEPTDGRKDVLYRALLHGI 300
DB 241 FILFCHNIKAPIYASGNLNIPFSTLIHKLGGFFIRRRIDEPTDGRKDVLYRALLHGI 300
QY 301 VELLRQOQFLEIFLEGTRSRGKTSACAGLLSVVDLTSTNVIPIIDLIIPVGISYDRII 360
DB 301 VELLRQOQFLEIFLEGTRSRGKTSACAGLLSVVDLTSTNVIPIIDLIIPVGISYDRII 360
QY 361 ECHYNGEQLKPKKNESLWSVARGVIRMLRNKVCVRVDFAPQPSKEYLESQKPVSA 420
DB 361 ECHYNGEQLKPKKNESLWSVARGVIRMLRNKVCVRVDFAPQPSKEYLESQKPVSA 420
QY 421 LLSLEQALLPAIILPSRPSDADEGRDTSINESRNATDESRLRLIANLAESHILFTASKSC 480
DB 421 LLSLEQALLPAIILPSRPSDADEGRDTSINESRNATDESRLRLIANLAESHILFTASKSC 480
QY 481 AIMSTHIVACLLLYRHQGDISTLVEDFFVMKEEVLARDFDLFGSGNSEDDVVMHAIQLL 540
DB 481 AIMSTHIVACLLLYRHQGDISTLVEDFFVMKEEVLARDFDLFGSGNSEDDVVMHAIQLL 540
QY 541 GNCVTIHTSRNDEFFITPTSPVPELNFYNGVLHVFMIEAIIACSLYAVLNKRGGLG 600
DB 541 GNCVTIHTSRNDEFFITPTSPVPELNFYNGVLHVFMIEAIIACSLYAVLNKRGGLG 600
QY 601 GPTSTPNLLISQBLVRKAASLCYLLSNEGTSISLPCQTFYQVCHETVKGFIQIGILTVAE 660
DB 601 GPTSTPNLLISQBLVRKAASLCYLLSNEGTSISLPCQTFYQVCHETVKGFIQIGILTVAE 660
QY 661 HDQOEDISPLAEOQMDKKLPELWSRSDDEDESDSDFGEQRDCYLKVSQKEHQPFITF 720
DB 661 HDQOEDISPLAEOQMDKKLPELWSRSDDEDESDSDFGEQRDCYLKVSQKEHQPFITF 720
QY 721 LQRLGLPLLAYSAAIIFVNFSPGVPPEPEYLOKLHKYILTRTRNVNAVVAESATYCLVK 780
DB 721 LQRLGLPLLAYSAAIIFVNFSPGVPPEPEYLOKLHKYILTRTRNVNAVVAESATYCLVK 780
QY 781 NAVMFKDIGVFKETKQKRVSVLELSSTFLPQCNQKLLLEYILSFVVL 828
DB 781 NAVMFKDIGVFKETKQKRVSVLELSSTFLPQCNQKLLLEYILSFVVL 828

RESULT 2
ID_PLSB_MOUSE STANDARD; PRT; 827 AA.
AC Q61586;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycerol-3-phosphate acyltransferase, mitochondrial precursor
```

```
QY 301 VELLRQOQFLEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPDILLIPVGSYDRII 360
Db 301 VELLRQOQFLEIFLEGTRSRSGKTSARAGVLSVVVNTLSSNTIPDILLIPVGSYDRII 360
QY 361 EGHYNGEOLGPKKNESLWSVARGVIRMLRKNYGVVVDFAQPFSLKYLESQSKPVSA 420
Db 361 EGHYNGEOLGPKKNESLWSVARGVIRMLRKNYGVVVDFAQPFSLKYLESQSKPVSA 420
QY 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSC 480
Db 421 PLSLEQALLPAILPSRPNDADEHQDLSNENSPADAFRRRLIANLAHILFTASKSC 480
QY 481 AIMSTHIVACLLYHRHQIDILSTLVEDFFVMKEEVLARDPDLGFGSGNSEDVWMAIQLL 540
Db 481 AIMSTHIVACLLYHRHQIHLSTLVEDFFVMKEEVLARDPDLGFGSGNSEDVWMAIQLL 540
QY 541 GNCVTITHTSRNDEFFITPTVPVSEFELNFSYNGVLHVFMEALIAACSLAVLNKRGIG 600
Db 541 GNCVTITHTSRNDEFFITPTVPVSEFELNFSYNGVLHVFMEALIAACSLAVLNKRGIG 600
QY 601 GPTSTPPNLSQEQVLVRAKASLCYLLSNEGTSISLPCQTFYQVCHETVGFQYIGLITVAE 660
Db 601 GSAGLGNLSQEQVLVRAKASLCYLLSNEGTSISLPCQTFYQVCHETVGFQYIGLITVAE 660
QY 661 HDDQEDISPLAEOQWKKLPEPLSWRDEDEDDSDGEGQRDCYLVKSQKEHQOQFITF 720
Db 661 QDDQEDVSPGLAEQWKKLPE-LNWRSDDEDEDDSDGEGQRDCYLVKSQKEHQOQFITF 719
QY 721 LQRLGLPILLEYSSAAIFVHNSGVPPEYLOKLHKLITRTENAVAYAESATYCLVK 780
Db 720 LQRLGLPILLEYSSAAIFVHNSGVPPEYLOKLHKLITRTENAVAYAESATYCLVK 779
QY 781 NAVMKFDIGVFKETQKRVSVLESLSTFLPCNRQKLLVILSPVVL 828
Db 780 NAVMKFDIGVFKETQKRVSVLESLSTFLPCNRQKLLVILSPVVL 827
RESULT 3
PLSB RAT STANDARD; PRT; 828 AA.
AC P97564; Q35349; P97565; P97566;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Glycerol-3-phosphate acyltransferase, mitochondrial precursor
DE (EC 2.3.1.15) (GPAT).
CN GPAM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND ACTIVITY.
RC TISSUE=Liver;
RX MEDLINE=99376617; PubMed=1046428;
RA Bhat B.G., Wang P., Kim J.-H., Black T.M., Lewin T.M.,
RA Fiedorek F.T. Jr., Coleman R.A.;
RT "Rat sn-glycerol-3-phosphate acyltransferase: molecular cloning and
RT characterization of the cDNA and expressed protein.";
RL Biochim. Biophys. Acta 1439:415-423(1999).
RN [2]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TOPOLOGY.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=20493538; PubMed=10924502;
RA Balija V.S., Chakraborty T.R., Nikonov A.V., Morimoto T., Haldar D.;
RT "Identification of two transmembrane regions and a cytosolic domain of
RT rat mitochondrial glycerophosphate acyltransferase.";
RL J. Biol. Chem. 275:31668-31673(2000).
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -!- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS. IT MAY
CC ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
```

```
outer membrane.
-!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF021348; AAB71605.1; -
DR EMBL; U36771; AAB39470.2; ALT INIT.
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
DR SMART; SM00563; Plsc; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Transmembrane; Mitochondrion; transit peptide.
FT TRANSIT 1 ?
FT CHAIN ? 828 MITOCHONDRION (POTENTIAL).
FT DOMAIN ? 471 GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
FT TRANSMEM 472 494 POTENTIAL.
FT DOMAIN 495 574 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 575 593 POTENTIAL.
FT DOMAIN 594 828 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT CONFLICT 37 37 C -> F (IN REF. 2).
FT CONFLICT 85 85 S -> P (IN REF. 2).
FT CONFLICT 300 300 I -> V (IN REF. 2).
FT CONFLICT 331 331 L -> V (IN REF. 2).
FT CONFLICT 472 474 ILF -> NLL (IN REF. 2).
FT CONFLICT 497 497 R -> W (IN REF. 2).
FT CONFLICT 602 606 SAGGL -> LPEP (IN REF. 2).
FT CONFLICT 644 644 Q -> H (IN REF. 2).
FT CONFLICT 744 744 G -> A (IN REF. 2).
SQ SEQUENCE 828 AA; 93714 MW; ACA4A087EBE12C CRC64;
Query Match 92.4%; Score 3955; DB 1; Length 828;
Best Local Similarity 91.7%; Pred. No. 2.1e-282;
Matches 759; Conservative 34; Mismatches 35; Indels 0; Gaps 0;
QY 1 MDESALTGTIDVSYLPHSSEYSVGRCKHTSEEWGECGRPTVFRSATLKWESLSMRKR 60
Db 1 MEESVVTGTTIDVSYLPNSSEYSYLGRCCKHTNEDWDCGKPTFRSATLKWESLSMRKR 60
QY 61 PFVGRCCVCTPQSQWDXKFNFSPISGLGRNVIYINETHRRHGLARLSYVLFQERDV 120
Db 61 PFVGRCCVCTPQSQWERFNFSPISGLGRNVIYINETHRRHGLARLSYVLFQERDV 120
QY 121 HKGMFATNVTENVLSRSRVQEAIAEVAELNPDGSAQQSKAVNKKKAKRILOEMVAT 180
Db 121 HKGMFATNVTENVLSRSRVQEAIAEVAELNPDGSAQQSKAIQKVKEKARKILOEMVAT 180
QY 181 VSPAMIRLTGWVLLKLFNSPFWNTQIHKGQLEWMAKATETNLLPLFVHRSHIDYLLLT 240
Db 181 VSPGMIRLTGWVLLKLFNSPFWNTQIHKGQLEWMAKATETNLLPLFVHRSHIDYLLLT 240
QY 241 FILFCHNIKAPYIAGNNLNPIESTLHKLGFFIRRLDETDPGRKDVYRALLHGH 300
Db 241 FILFCHNIKAPYIAGNNLNPIESTLHKLGFFIRRLDETDPGRKDVYRALLHGH 300
QY 301 VELLRQOQFLEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPDILLIPVGSYDRII 360
Db 301 VELLRQOQFLEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPDILLIPVGSYDRII 360
QY 361 EGHYNGEOLGPKKNESLWSVARGVIRMLRKNYGVVVDFAQPFSLKYLESQSKPVSA 420
Db 361 EGHYNGEOLGPKKNESLWSVARGVIRMLRKNYGVVVDFAQPFSLKYLESQSKPVSA 420
QY 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSC 480
Db 421 PLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSC 480
QY 481 AIMSTHIVACLLYHRHQIDILSTLVEDFFVMKEEVLARDPDLGFGSGNSEDVWMAIQLL 540
```

Db 481 AINMTHIVACLLYRHRQGIHLSTLVEDFPVMKEEVLARDFDLFGSGNSDDVVMHAIQLL 540  
QY 541 GNCVTHIHSRNDDEFTPTPTVPSFELNFGVNLHVEIMAEIACSLIYAVLNKRGIG 600  
Db 541 GNCVTHIHSRNDDEFTPTPTVPSFELNFGVNLHVEIMAEIACSLIYAVLNKRGIG 600  
QY 601 GPTSTPNLISQOLVRKAASLCVLLSNEGTISLPQTFYQVCHETVGRFIOYGLITVAE 660  
Db 601 GSAGGLNLISQOLVRKAASLCVLLSNEGTISLPQTFYQVCHETVGRFIOYGLITVAE 660  
QY 661 HDQEDISPLARQOMDKKLPELWSRDEDESDFGEEQRDCYLKVSQSKHEHQFIF 720  
Db 661 QDQEDVSPGLARQOMDKKLPELWSRDEDESDFGEEQRDCYLKVSQSKHEHQFIF 720  
QY 721 LQRLGLLEAYSSAIFVNFSPGVPPEPYLOKLHLYLTRNNAVVAESATYCLVK 780  
Db 721 LQRLGLLEAYSSAIFVNFSPGVPPEPYLOKLHLYLTRNNAVVAESATYCLVK 780  
QY 781 NAVMFKDIGVFKETKQKRVSVLELSTFLPQCNROKLEVIILSFVVL 828  
Db 781 NAVMFKDIGVFKETKQKRVSVLELSTFLPQCNROKLEVIILSFVVL 828

## RESULT 4

PLSB CAEL STANDARD; PRT; 718 AA.  
ID PLSB CAEL  
AC Q22949;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Probable glycerol-3-phosphate acyltransferase, mitochondrial precursor  
DE (EC 2.3.1.15) (GPAT).  
GN F08F3.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Blanchard M., Bradshaw H.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-  
CC acyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS. IT MAY  
CC ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
CC (Potential).  
CC -1- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U64847; AAB04876.1; -.  
DR PIR; T29448; T29448.  
DR WormPep; F08F3.2; CE09258.  
DR InterPro; IPR002123; Acyltransferase.  
DR Pfam; PFO1553; Acyltransferase; 1.  
DR SMART; SM00563; Pfam; 1.  
KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
KW Transmembrane; Mitochondrion; Transit peptide.  
FT TRANSIT 1 ?  
FT CHAIN ? 718 MITOCHONDRION (POTENTIAL).  
FT PROBABLE GLYCEROL-3-PHOSPHATE  
FT ACYLTRANSFERASE.  
FT TRANSMEM 409 425  
FT POTENTIAL.  
SQ SEQUENCE 718 AA; 82071 MW; E0A36A4A86FC138D CRC64;

Query Match 15.8%; Score 678; DB 1; Length 718;  
Best Local Similarity 27.2%; Pred. No. 6.8e-42;  
Matches 197; Conservative 148; Mismatches 264; Indels 114; Gaps 21;  
QY 128 NTVNVLNRSRVOEAEVAELNPDGSAQOQSKAVNKVKKAKRILQEWVATVSPAMIR 187  
Db 78 NVEKDVLSKRVRHVISKL-----KEQN---DEQNRVQFTEISARLSFKICK 124  
QY 188 LTGWLLKLFNFFNFIQIHKGQLEMVKAATNPLPLFLPVHRSHIDYLLTFLFCHN 247  
Db 125 CCSYVLVYKFRMLMDKLLVCKEEMVLYEAGTGPVYVLPVLRSHDYLITTCNWHFG 184  
QY 248 IKAPYIASGNLNIPIFTLIHKLGGFFIRRLDETDPGRKDVLYRALLHGHIVELLROQ 307  
Db 185 LXLPHIASGDNLSLGLWLLRATGATGAFRRVDPDERGKQLYRALLHSYIEQVLSKD 244  
QY 308 QFLEIPLGTRSRSGKTS CARAGLLSVVVDITLSTNVIPDILIPVGISVDRIIEGHYNGE 367  
Db 245 MPTEFLEGTGRSRFGKALTPKNGLISNVVVEAVQHGFICDCYLPVPSYTYDAVVEGIFLHE 304  
QY 368 QLGKPKNESLMSVARGVIRMLRKNYGC--VRVDFAPQFSLKEYLESQSKPVSALLSLE 425  
Db 305 LMGIPIKVRSEVLGVRGIFSGFSKQCGVVRMHYGRPIRLTEYLAT-----ITASLSSN 359  
QY 426 QALLPAILPSRPSDAADGRDTSINES-----RNATDESRRRLIANIAEHILFT 475  
Db 360 HRTRPV-----RMTKLSFSYRELVPWHRTHSETVDDRTWIRALGFHVVE 406  
QY 476 ASKSCAIMSTHIVACLLYRHRQGIHLSTLVBDFVFMKEEVLARDPD-LGFSG---NSED 531  
Db 407 AQMWCISPVAVVSCILLAKWRGKVSRTFERDCWLCCKIIAEGGDVVGYOSKTKGSA 466  
QY 532 VVMHAIQLGNCVTIHTSRNDEFFITPTTVPSPFELNFGVNLHVEIMAEIACSLY 591  
Db 467 LVKYAFKLEKSEVEV-----DE-VSPKSHSFTITLAYNKNVTCRFSIKSVIALTI- 519  
QY 592 AVLNRKGLGGTSTTPPNLISQOLVRKAASLCVLLSNEGTISLPQTFYQVCHETVCKFI 651  
Db 520 -----VSRPSGTK---LSIQIVEDALSCLDWLQFEPFCPCDLSRELNVHLG--- 566  
QY 652 QYGILTVAEHDQEDISPLARQOMDKKLPELWSRDEDESDFGEEQRDCYLK 708  
Db 567 -----OKEWS-----DPIHGFRLSEIEDDGFLDAGGALNSGTLRV 601  
QY 709 SOSKEHQQITFLORLLGLLEAYSSAIFVNFSPGVPPEPYLOKLHLYLTR----- 763  
Db 602 RDAKS-RETLPFANLVRPFVQSILYLISSFVSEKCP-EPSTSDNNIIRLOCOQSLAGDI 659  
QY 764 ERNAVVAESATYCLVKNVAKMFKDIGVFKETKQKRV-----VLELSSTFLPQCNROK 817  
Db 660 DLPPAPLESINSDSFKNALRVLKDKGLLQRTSPNSTARSGNSRLAELIS-----NLER 713  
QY 818 LLE 820  
Db 714 VLE 716  
RESULT 5  
PLSB XYLFA STANDARD; PRT; 870 AA.  
ID PLSB XYLFA  
AC Q9PEJ37;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).  
GN PLSB OR XF1031.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xylella.  
OX NCBI\_TaxID=23711;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9a5c;

RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Canagar L.E.A., Carraro D.M., Carrier H.,  
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neco C.M.,  
 RA Coutinho L.P., Cristofani M., Dias-Neco E., Docena C., El-Dorri H.,  
 RA Faciniani L.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kurama E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.R., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Mend C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.B. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
 RA "The genome sequence of the plant pathogen Xylella fastidiosa";  
 RT Nature 406:151-159(2000).  
 CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-  
 CC acyl-sn-glycerol 3-phosphate.  
 CC -!- PATHWAY: DE NOVO PHOSPHOLIPID BIOSYNTHESIS; FIRST STEP. MAY ALSO  
 CC FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE003940; AAF83841.1; -  
 CC PIR; B82732; B82732.  
 CC HAMAP; MF\_00393; -; 1.  
 CC InterPro; IPR002123; Acyltransferase.  
 CC Pfam; PF01553; Acyltransferase; 1.  
 CC SMART; SM00563; PlsC; 1.  
 CC Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;  
 CC Complete proteome.  
 CC SEQUENCE 870 AA; 98483 MW; 52A8ADA9DFB9868 CRC64;  
 SQ  
 Query Match 11.8%; Score 505; DB 1; Length 870;  
 Best Local Similarity 22.4%; Pred. No. 4.5e-29;  
 Matches 197; Conservative 148; Mismatches 314; Indels 222; Gaps 26;  
 QY 13 VSYLPHSEYSVGR-CKHTSEWGC-----GPRPTVFSATLKWESLMSRKPVRVG 64  
 DB 64 VOYPEHCVOYNDGRIIYVLEDYGLCNTLILDKACRKTLPSPILFPGNLPQRKRAYLA 123  
 QY 65 RCVCYCT-----PQSWDFKFNPSIPSL-----GLRNV-----IYNETHTRHGWLA 106  
 DB 124 LSRRSSNSLIPNQRGGKTHSDSLANLQAHRIITDLVHLVPVSIFIGTRPDROSGWFA 183  
 QY 107 -----RLSYVLPIQ 116  
 DB 184 VLFSENWALVGRFRLAVLLNGRNTIVCFAPPISVRQTLNEGLPPERTLRKQLVRAH 243  
 QY 117 ERDVHKGHFAINV-----ENVLSRSVQEAIAEVAALPDGSAQQSQKAVNKVKKA 170  
 DB 244 FRIRRETIVGPDLSRLLVDVNLATEAVREAIA-----SQAKRDGTDLSLSTWKA 294

QY 171 KRILQEWATVSPAMIRLTGWLLKLKLFNSFWNIQIHKGQLEMKAAATETNLPFLPVP 230  
 DB 295 QAYAEIADYSSPVIRSADFLFHVWNRITAGVLIH--HVDSEK-ETAPGHEVVVPSH 351  
 QY 231 RSHIDYLLLTFLFCHNIKAPYIASGNLNIPSTLIHKLGFGFPFIRRLDETDPGRKDV 290  
 DB 352 RSHIDYLLSLCYQCQGVLPFHIVAGINLNLPIVGTLLRKCCAFIRRSI-----KGNM 405  
 QY 291 LYRALLGHIVELARQQQFLIELEGTSRSGKTSARAGLLSVVVDLTSTNVIPDILII 350  
 DB 406 LYSIVLSYVAQLVAGGYSLEYFIEGSRSTRGLLPQKGGIMMTLQAFLRQPRRPVLFQ 465  
 QY 351 PVGISYDRIIEG-HYNGEQLGPKKNESLWSVARGVIRMLRKNYGCVRVDVDAQPESLKEY 409  
 DB 466 PIYIGYEKLIBGTSYLDSECEPKKESIWELFWNIPKVLQKYGQVVVNFGEPIALNDV 525  
 QY 410 LESQSKPVSALESQALLPAIILPSRPSDAADSGRDTSTINESRNATDESURRLIANIA 469  
 DB 526 L-----AELAPEWEGQALNENEKPAMLS---NTVNHIA 555  
 QY 470 EHILFTASKSCAIMSTHIVACILLYRHQGDIDLTSLVEDFFVMKEEVLARDDFLGFSG-- 527  
 DB 556 ROIQTRINSAADVNPINILALALISTPKHAGEADLIAQITLCKKILL-----EUPYSNRV 611  
 QY 528 ----NSEDVVMHAIQLIGNCVITITHTSRNDEFFITPTTVPVSFELNFYNGVLHVFM 582  
 DB 612 TVTPTHTPERIAHAEQI--NILTRVHHPLGDLVLDGDNV-----LLSYFNNVLHLFTA 665  
 QY 583 BAITACSLYAVLNKRGGLGPTTPNLIISQOLVRKAASLCYLLSNETGISLPQOTFYQV 642  
 DB 666 SAWVACCFKN--NRR-----ISRIALRLGVMPFLQAEFLPWTEDQFAQH 711  
 QY 643 CHETVGKFIQVILTVAEHDQEDISPSLAQOWDKLPELWSRSEDESDSGESQR 702  
 DB 712 IQQVIELFVREGLL-LSAGDEED-----PLTRNTSQTDE----- 745  
 QY 703 DCYLKVSQSKHQFIPTFLQRLGLLEA-----YSSAAIFVHNFGSPVPEYLQKLHK 757  
 DB 746 -----VFLRAISHSQAFERYITISILVKNPGFTLSASE-----LES 785  
 QY 758 YLITRTNRNAVYAESATYCLVKNVAKMFKDIGVFKETKQK 798  
 DB 786 LCQLAAQRLSLLYASTA-----PEFPDKGLFRGFIQK 817  
 RESULT 6  
 ID PLSB\_PSEAE STANDARD; PRT; 834 AA.  
 AC Q9HXW7;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).  
 GN PLSB OR PA3673.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 CC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 15692 / PAOI;  
 RA MEDLINE=2043737; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RA "Complete genome sequence of Pseudomonas aeruginosa PAOI, an  
 RT opportunistic pathogen."; Nature 406:959-964(2000).  
 RL Nature 406:959-964(2000).  
 CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-



Qy	107	RLSLVLFQBRDHHKGMFATNVT-----ENVLSRSRVOEATAEVAEELNPDGSAQQOS	160
Db	245	RKLQRLVLFTHRRREAVIGPDLSSTRELLVDVQLAAESVREATAIAQAKDN-----S	296
Qy	161	KAVNKVKKKAKRIQLQEMVATVSPAMIRLTGWLKLLFNSPFNNIQLHKQLEMLKAAATET	220
Db	297	KPVD--AWRKAHAYAEATAADYSSPVRSASFLTLTHVMNRIYAGVLVH--HLDKLQKAAPG	353
Qy	221	NLPILFLFVHRSHDYLLNLFILFCHNLIKAPYATASGNINIPTESTLHKLGCFEFTTRRL	280

SEQUENCE FROM N.A.  
RC STRAIN=306 / ATCC 13902 / XV 101;  
RX MEDLINE=22022145; PubMed=12042417;  
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
RA Camarotte G., Cannavan F., Cardoso J., Chambergro F., Clapina L.P.,  
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,  
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
RA Locali A.C., Machado L.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
RA Moreira E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
RA Moreira L.M., Novo M.T.M., Okura V.C., Oliveira M.C., Oliveira V.R.,  
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
RA Spínola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.P.,  
RA Setubal J.C., Kitzajma J.P.;  
RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
RT host specificities.";  
RL Nature 417:459-463 (2002).





```
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
DR SMART; SM00563; PlsC; 1.
KW Transferase; Acyltransferase; Peroxisome; Membrane.
FT DOMAIN 5 POLY-SER.
FT SITE 676 678 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 678 AA; 76869 MW; FA7245A2DDDD174FF CRC64;

Query Match
Best Local Similarity 9.9%; Score 424; DB 1; Length 678;
Matches 166; Conservative 122; Mismatches 255; Indels 148; Gaps 23;

QY 133 VLNSRVQEAIAEVAELNPDGSAQQOQKAVKVKKAKRILOEMVATVSPAMRLTGWV 192
DB 74 VLSSEINVIKLGRE-----SLTGVDVLEEEASEILEENSHKLRIGAIRFAFF 124
QY 193 LKLFNSFFWNTQIHKGLEVMKATETNLPLLELPVHRSHDYLLFLFLFCHNIKAPY 252
DB 125 LSKIFQIFSKVCVNEEGIQKQRAVQEH-PVLLPSHRSYIDFLMLSFILYSYDLPPV 183
QY 253 IASGNN-LNIPFSTLIHKLGGFFIRRLDETDPGRKQVLYRALLGHIVELLR-QQOFL 310
DB 184 IAAGMDFLGMVVSELLRMSGAFFRR---TFGNGK-LYWAVSEYVTKMLRCGYAPV 237
QY 311 EFLFEGTSSRSKTSARAGLLSVVDVLTSTNVIPIILIPVIGISYDRIL-EGHYNGQL 369
DB 238 EFLFEGTSSRAAKLTTPKFLGLNIVMEPFKREVDYTFVPITISYDKILBESLYAYEIL 297
QY 370 GPKPNESLWSVARGVIRMLKNGCVVDVFAQPESLKEYLESQKVPVSALLSLEQALL 429
DB 298 GVPKPEKSTTGLK-ARRILSENFGSIHVYFGDVPVSLRSLAAGRLNRNT-----YNLV 349
QY 430 PAILPSR-PSDAADGRTDSINESRNATDESRLRL--IANLAHILFTASKSCAIMTSH 486
DB 350 PRCIQKQPEDV-----QAFVTEVAYKMLQIQENLA-----LSPW 385
QY 487 IVACLLLYRHRQIDLTSLVEDFVYKBEVLARDFDLGFSGNS--EDVVMHAIQLLNCV 544
DB 386 LLVVTILLQNLQSLDFDALVEKTLWLGVTQVFGGFLLPDKNKLPVEVVQSSILLHSNLA 445
QY 545 TITHSRNDEFFITST-----TVPSVEL---NFSYNGVLHVFMELIACSLYAV 593
DB 446 SLV-----KQVVVLKNGSSQVVGVLVPEHTIALMCSAYRNQLNIFARPISLVALAHMT 501
QY 594 -LNKRGGLGPTSTPPNLIISQBLVRKAASL-----CYLLSNEGTISLPCQTFYQVCH 645
DB 502 PGLREDVFSCFSFLRNVSDEFILPGNTLRDPFEGCYLLCKAAMQV-----550
QY 646 TVGKFIQYIGILTVAEHDDQEDISPLAQQWKKLPEPLSMRSDDEDSDFGEQRDCY 705
DB 551 -AGKDI---ILT-----DK-----560
QY 706 LKVSQKHEHQOITFLQRLGLLLEAYSAALFVNFSGVPPEPYLOKLHKYLITRTER 765
DB 561 -----GTAVLOFLTSLFKPFVSQYLLCHRYLLHEEDYFGEKEYLAARKFTQLLDQ 612
QY 766 NVAVVAESATYCLVKNVAKMKDGVKFKTK 796
DB 613 GSSQCYDALSELQKNLAFAFVRLGVVKKK 643

RESULT 11
ID_PLSB_VIBVU STANDARD; PRT; 809 AA.
AC Q8DD48;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).
GN PLSB OR VV1165.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
```

```
NCBI_TaxID=672;
[1]
SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H.; Kim S.Y.; Chung S.S.; Kim J.J.; Moon Y.H.; Jeong H.;
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -!- PATHWAY: De novo phospholipid biosynthesis; first step. May also
CC function in the regulation of membrane biogenesis.
CC -!- SUBCELLULAR LOCATION: Membrane-Bound (By similarity).
CC -!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB016800; AAC09636.1; -
DR HAMAP; MF_00393; -; 1.
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
DR SMART; SM00563; PlsC; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;
KW Complete proteome.
SQ SEQUENCE 809 AA; 90596 MW; 0E04FF0B4980DCDC CRC64;

Query Match
Best Local Similarity 9.3%; Score 399.5; DB 1; Length 809;
Matches 180; Conservative 145; Mismatches 223; Indels 257; Gaps 38;

QY 84 PSLGLRNVIIYNETH-----TRHRGWLARRLSYVLFTEQERDVHKG-----MFATNVT 130
DB 169 PVVSLR---YMAQSHGTDSSIAHK--LARVATHFSRQKLAASGPDLPSSRQVLF-- 218
QY 131 ENVLNSRVQEAIAEVAELNPDGSAQQOQKAVKVKKAKRILOEMVATVSPAMI----- 186
DB 219 -RLMKSPAIEQAIEE-----EAKNKNISMEKARKEAODIMDEIAADPSYSLVKQGD 268
QY 187 RLTCGWLLKLLFNSFFWNTQIHKGLEVMKATETNL-----PLFLPVRSHDYLLLTFF 241
DB 269 RLLGWL-----WN-KLYQG-LNINNAATVRRLAQDGHGHEIVIVPCHRSHMDYLLSY 317
QY 242 ILFCHNIKAPYIASGNNLNI-----PIFSTLIHKLGGFFIRRLDETDPGRKQVLYRALLH 297
DB 318 VLXHEGVVPHIAAGINLFPFAGPIF---RRGGAFFIRSF-----KGNRLYSTIFR 367
QY 298 GHIVELLRQQOFLBEIFLEGTSRSKTSARAGLLSVVDVLTSTNVIPIILIPVIGISYD 357
DB 368 EYLAELFAKGYSEVYFSGGSRRTGRLLPAKTGMLAMTIQAMLRGLNRPVTLVPVYIGE 427
QY 358 RIIE-GHYNGEQLGPKPKKNESLWSVARGVIRMLK--NYGCVRVDFAPQPSLKEYLESQS 414
DB 428 HVMEMATYAKELRGKRKEKEN---AGLVRLTKLRNFGLVYGFGEPIPLNQVLLNSHA 483
QY 415 QK-----PVSALLSLEQALLPAILPSRPSDAADGRTDSINESRNATDESRLRLIANL 468
DB 484 PEWTKDIDPMCA-----SRPQ-----WINP-----VVNQL 508
QY 469 AEHLFTASKSCAIMSTHIVACLLYRHRQIDLTSLVED---FFVVKERVLARDFDLG 524
DB 509 ANKMWTHINDAAAAANALTLCATALLASQRALSXDSLHQIECYLQLLKNVPYSKTYTPV 568
QY 525 FSGNSDVMVHAIOL-----LGNCVITHTSRNDEFFITSTVPSVFLNFSNG 575
DB 569 -SESAAALVEHAISLDKFVETDTMGDIISL---DRNQSLMT-----YRNN 612
QY 576 VLHVFIEMAIACSLYAVLNKRLGGTSTTPPNLISQBLVRKAASLYLLSNEGTSIP 635
```

```
Db 613 IHLFALPSLIA-----QMIIRGEN-----632
QY 636 CQTFYQVCHETVGFIOYIGILTVAEHDQD-EDISPSLAEOQWKKLPEPLSWRSDEDE-693
Db 633 -----LTVSQQQVAEIEYFPLKAEFL-----LSHKEEELDEL 665
QY 694 -----DSDGFEORDCVLKSQSKHEQQTFLQRLGLPLLEAYSSAAIFVHNFSG 744
Db 666 VVKVNLVSQDLISLKEB---KVAKQANTLVLVGLRITSETLQRYSTA-----FNL 716
QY 745 PVPEPEYLOKLHKYLIITRTERNAVVAESATYCLVKNVAVKDFDGVFKETQKRVSVLE 804
Db 717 LVSNPE---LAKADLEQKSDI---AQLRLHGINAPEYF-DKGVF-----AS 758
QY 805 LSSTFLPDQ-----CNROKLLLEY 821
Db 759 LFSTLKOQGYLDSDGNCDEKTAQF 783

RESULT 12
PLSB_VIBPA STANDARD; PRT; 808 AA.
AC Q87KNO;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).
GN PLSB OR VP2947.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshina K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RL distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
CC -I- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -I- PATHWAY: De novo phospholipid biosynthesis; first step. May also
CC function in the regulation of membrane biogenesis.
CC -I- SUBCELLULAR LOCATION: Membrane-bound (By similarity).
CC -I- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; APO05083; BAC61210.1; .
DR HAMAP; MF 00393; .
KW Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;
KW Complete proteome.
SQ SEQUENCE 808 AA; 91143 MW; 96E9BAEA6DCE8B61 CRC64;

Query Match 9.3%; Score 398.5; DB 1; Length 808;
Best Local Similarity 20.9%; Pred. No. 2.7e-21;
Matches 160; Conservative 144; Mismatches 264; Indels 197; Gaps 29;

QY 84 PSLGLRNVYINETH-----TRHGWLARLSYLVLFQERDVHKG-----MFATNVT 130
Db 169 PVVSLR---YNADSHGTDISAHK---LARVARIHFSRQKLAASGPNLPQRQVLFA---- 218
QY 131 ENVLNRSRVQEAIAEVAELNPDGSAQQQSKAVNKVKKAKRIIQEMVATVSPAMI---- 186
```

```
Db 219 -RLKSPATEQAI-----EDEAKSKDISIEKARKEAHDMDEIAADPSVGLVKNGD 268
QY 187 RLTCGWLLKLPNSFPWNIOIHKGQLEWKAATETNLPFLPVRHSHIDYLLITFLFCH 246
Db 269 RILSMLTKLYOG-----LHINNASTVRRLAQDGHEIVVVPCHRSMDYLLSLLSYLH 322
QY 247 NIKAPYIASGNLNI-----PIFSTLIHKLGFFIRRLDETDPGRKDVLYRALLHGHIVE 302
Db 323 GNVPPHIAAGININFPAGPIF-----RRGGAFFIRSP-----KGNK--LYSTIFREYLA 372
QY 303 LLRQOQFLBIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPDILIIIPVGISYDRII 361
Db 373 LPAKGSYVEVSEGGSRGRTGRLLQAKTGLAMTIQAMLRLGLNRPVTLVPVYIGYEHMEV 432
QY 362 GHYNGEOLKPKKNESLWSVARGVIRMLRK--NYGCVRVDPFAQPPSLKYLSELSQSKPVS 419
Db 433 GTYAKELRGKRKEKEN-----AGLVLTLRKLNFGLYVNFGEPIQLNLYNEHA----- 483
QY 420 ALLSLEQALLPAILSPRPSDAADGRTSINERSNATDESRRRLIANIAELHILPTASKS 479
Db 484 -----PEWTXDI-----DSMGSKPQWNPVNVNELANKMTHINDA 519
QY 480 CAIMSTHIVACLLLYRHRQGDILSTLVED-----FFVMKEEVLARDFDLFGSGNSEDVVMH 535
Db 520 AAANALTLATALLASRQALSRLDSLINQIECVLKLKNNPYSSSTSTIP-TESAELVDH 578
QY 536 AIQL-----LGNCVTIHTSRNDEFFITPSTVPFELNFSYNGVLHVHFMELAI 586
Db 579 AISLDKFVETDSMGDIISL---DRSQSILMT-----YRNNIIHIFALPSLI 623
QY 587 ACSLYAVLNKRGIGGPTSTPPNLIQEOQLVRKAASLCYLLSNEGTLISLPCOTFYQVCHET 646
Db 624 A---QMIIRQRL-----TVEKIQENVAQIYPLPKELFLSYOEEDLNDLVVKT 669
QY 647 VGKFIQYIGIL-----TVAEHDDQEDISPSLAEOQWKKLPEPL 684
Db 670 LNEFAEQMKICLDGNKLEINQSNQNPVLVLLGRTITETLQRYSIAMNLL-----VAYPE 722
QY 685 SWRSDEBDDSDFGEEQRDCYLVKSQSKHQQTFFLQRLGLPLLEAYSSAIFAIVNFSG 744
Db 723 LKGSLEQKSDIAQR-----LGRHGINAPEFPDFKGVFTAMFN- 761
QY 745 PVPEPEYL-----QKLHKYLIITRTERNAVVAESATVCL 778
Db 762 TLKQBYLSDGNCCKKTKQKAKLLFTLLYPEVKLTIESIHQL 806

RESULT 13
PLSB_SALTI STANDARD; PRT; 806 AA.
AC Q8ZIF6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).
GN PLSB OR STY4431 OR T4141.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feitwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
```



Db 211 PARQDLNFKLLAKAIARAVED-----EARSKKISHEKAQQAQNAIALMBEEIAANSFYEM 263  
Qy 186 IRLTGWLLKLFNSFFWNIOIHKQLEWKAATETNPLFLPVHRSHIDYLLTLTFLC 245  
Db 264 IRLTDLGLTGNKLYOGINVHNA--BRVQLAHGHEIIVVPCHRSHMDYLLLSVLYH 321  
Qy 246 HNIKAPYIASGNLNI-----PIFSTLIHKLGGFIRRRRLDETDPGRDVLRYALLGHIV 301  
Db 322 QGLVPPHIAAGINLFWPAGPIP-----RLGAPFIR-----TFKGNK--LYSTVFREYLG 371  
Qy 302 ELLRQOQFLIELEGTSRSGKTS CARAGLLSVVVDLTNNVPDILIIIPVGSYDRIIE 361  
Db 372 ELFSRGYSVEYFVEGGSRGRTGLLDPKTGTLSMTIOAMLRGTRPITLPIYIGYHME 431  
Qy 362 -GHVNGQLKPKKNESLWSVARVIRMLKNGYCVRVDFQAQPSLKEYLES-----QS 414  
Db 432 VGTAKELRGATKEKSLPQMLKSLK--RLNGQGYVNFGEPMPLTYLNNQHVPEWRES 489  
Qy 415 QKPVSAALLSLEQALLPA--ILPSRPSDAADSGRDTSSINESRNATDESRLRLIANLAHIL 473  
Db 490 IDPI-----EAIKPAWLPTVNSIAAD-----LM 513  
Qy 474 FTASKCAIMSTHIVACLLYHRHQGIDLTSLVED-----FVWKEVLEARDPDLGFSGNS 529  
Db 514 VRINNAGANAMNLCTALLASRQSLTREQLTEQLDCYLDLNRNVSTSTVP--AASA 572  
Qy 530 EDVVMHAIOLGNCVTITHTSRNDEFFITPSTVPSVFNLFYNGVLHVFMIEAIACS 589  
Db 573 GELIAHALQW--NKFEVEKDTIGDIILPREQAV-----LMTYRNNIAHMLPESLMAAI 626  
Qy 590 LYAVLNKRLGGTSTPPNLIISOBLVRKAASLCYLLSNEGTSILPCQTYQVCHETVGK 649  
Db 627 I-----TQHRISRDALQOHVEALYPMLKAEFLRWEREELASVIDALASE 672  
Qy 650 FIOVGLITVAEHDQEDISPS 670  
Db 673 MQRQGLITL--QDELHINPT 691

## RESULT 15

PLSB VIBCH STANDARD; PRT; 811 AA.  
ID PLSB VIBCH  
AC Q9KPB;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glycerol-3-phosphate acyltransferase (EC 2.3.1.15) (GPAT).  
GN PLSB OR VC0093.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=E1 Tor N16961 / Serotype O1;  
RX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
RT cholerae.";  
RL Nature 406:477-483(2000).  
CC -I- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-  
CC acyl-sn-glycerol 3-phosphate.  
CC -I- PATHWAY: DE NOVO PHOSPHOLIPID BIOSYNTHESIS; FIRST STEP. MAY ALSO  
CC -I- SUBCELLULAR LOCATION: Membrane-bound (by similarity).  
CC -I- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE004100; AAF93271.1; --  
DR TIGR; B82365; B82365.  
DR HAMAP; MF\_00393; --; 1.  
DR InterPro; IPR002123; Acyltransferase.  
DR Pfam; PF01553; Acyltransferase; 1.  
DR SMART; SM00563; Plsc; 1  
KW Phospholipid biosynthesis; Transferase; Acyltransferase; Membrane;  
KW Complete proteome.  
SQ SEQUENCE 811 AA; 90824 MW; B076969A30BCFC8D CRC64;

Query Match 9.3%; Score 396; DB 1; Length 811;  
Best Local Similarity 23.3%; Pred. No. 4.1e-21;  
Matches 157; Conservative 119; Mismatches 217; Indels 182; Gaps 30;

Qy 47 ATLKW--KESLSMRKRPV-----GRCCYCTPQSDKFPNPSISLGLR 89  
Db 129 ATVLWGRKPKGQBRPYLQALNGPEKALAVLASGRDCL-----VRFSPV----- 174  
Qy 90 NVIYNETH-----TTRGWLARRLSYVLFIOERDVHK-----MFATNTVENVLNS 136  
Db 175 SMRYADHTTGDASTAHK--LAVARIHFSRQKLAASGNLPQRAQLFA-----RLMNS 226  
Qy 137 SRVOEAIKVAELNPDGSAQOQSKAVNKVKKAKRILEMVAIVSPAMI-----RLTGWV 192  
Db 227 PAIEKAIAD-----EAKSKQIPLEKARKEHDLDEIAADFSYLVKKGDRILGWL 277  
Qy 193 LKLFNSFFWNIQIHKQLEWKAATETNL-----PLFLPVHRSHIDYLLTLTFLC 247  
Db 278 -----WN-RIVOG-LNINNAATVRLAQDGHIEIVVPCHRSHMDYLLLSVLYHEG 326  
Qy 248 IKAPYIASGNLNI-----PIFSTLIHKLGGFIRRRRLDETDPGRDVLRYALLGHIVEL 303  
Db 327 MVPPHIAAGINLFWPAGPIP-----RRGAPFIRRSFKGAP-----LYSTIFREYLAEL 376  
Qy 304 LRQOQFLIEFLGTSRSGKTS CARAGLLSVVVDLTNNVPDILIIIPVGSYDRIIE-G 362  
Db 377 FAKGYSVEYFVEGGSRGRTGLLPKATGLMAMTIOAMLRGNRPVTLVPVYIGYHVEVG 436  
Qy 363 HYNGEOLGPKKNESLWSVARVIRMLK--NYGCVRVDFQAQPSLKEYLESQSKPVSA 420  
Db 437 TYAKELRGKKEKEN-----AGLVRLTLKRLNFGQGVNFGFPIPLNQFLNE----- 484  
Qy 421 LLSLEQALLPAILPSRPSDAADSGRDTSSINESRNATDESRLRLIANLAHILFTASKSC 480  
Db 485 -----TVPOWTQDIDPMGE--SKPQMTPTVKNLANRMMTHINDAAVNAATLTC 531  
Qy 481 AIMSTHIVACLLYHRHQGIDLTSLVEDFFVWKEV--LARDFDLGF-----SGNSEDV 532  
Db 532 A-----TALLASRQ-----ALARDNLIKQVDCYLLNRNVPSYATSTLPSEAEKL 578  
Qy 533 VMHAIQL-----LGNCVTITHTSRNDEFFITPSTVPSVFNLFYNGVLHVFMIE 583  
Db 579 VEHAESLDKFFVETDMGDIISL--DRNQSIILMT-----YRNNIHLALP 623  
Qy 584 ATIACSLYAVLNKRLGGTSTPPNLIISOBLVRKAASLCYLLSNEGTSILPCQTYQVVC 643  
Db 624 SLIAQLLIR-----QQSVLSLEKQVATVAQIYPFELKQELFLRFAEELNDLV 669  
Qy 644 HETVKGFIQYGLTV 658  
Db 670 LRYVAELARQGLTV 684

Search completed: January 10, 2004, 20:19:17

Job time : 29 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2004, 19:55:42 ; Search time 35 Seconds  
(without alignments)  
2275.076 Million cell updates/sec

Title: US-09-935-290-2  
Perfect score: 4280  
Sequence: 1 MDESALTGTIDVSYLPHSS.....FLPQCNRKLLLEYLSFVVL 828  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3985.5	93.1	827	A41672	glycerol-3-phospha
2	678	15.8	718	T29448	hypothetical prote
3	505	11.8	870	B82732	glycerol-3-phospha
4	462	10.8	834	F83185	glycerol-3-phospha
5	397	9.3	806	AB1015	glycerol-3-phospha
6	396	9.3	811	B82365	glycerol-3-phospha
7	395	9.2	810	D64090	glycerol-3-phospha
8	391.5	9.1	827	XURCAG	glycerol-3-phospha
9	391	9.1	825	AC0039	glycerol-3-phospha
10	390.5	9.1	827	H91256	glycerol-3-phospha
11	387.5	9.1	827	D86097	glycerol-3-phospha
12	289	6.8	775	T45238	probable transfera
13	245	5.7	789	A70868	probable plsb2 - M
14	229	5.4	621	D70762	probable transfera
15	150	3.5	287	AG2414	hypothetical prote
16	140	3.3	3144	A46068	Huntington disease
17	135.5	3.2	239	F96984	1-acyl-sn-glycerol
18	134	3.1	240	S75162	hypothetical prote
19	128.5	3.0	967	T40171	probable chromosom
20	124.5	2.9	1176	A27826	DNA-directed RNA p
21	119.5	2.8	1174	S57060	probable membrane
22	118	2.8	1258	T30252	nuclear protein SA
23	116	2.7	1262	T30524	protein phosphatas
24	115.5	2.7	3119	I49729	HD protein - mouse
25	115	2.7	3744	S45715	hypothetical prote
26	111	2.6	829	F89630	protein F59F5.7 [i
27	111	2.6	987	1 JQ0016	6-phosphofructokin
28	110.5	2.6	820	D71471	probable DNA misma
29	110	2.6	651	T15624	hypothetical prote

RESULT 1

A41672

glycerol-3-phosphate acyltransferase homolog - mouse  
C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 05-Nov-1999

C:Accession: A41672

R.Shin, D.H.; Paulauskis, J.D.; Moustaid, N.; Sul, H.S.

J. Biol. Chem. 266, 23834-23839, 1991

A:Title: Transcriptional regulation of p90 with sequence homology to Escherichia coli gl

A:Reference number: A41672; MUID:92084678; PMID:1721057

A:Accession: A41672

A:Molecule type: mRNA

A:Residues: 1-827 <SHI>

A:Cross-references: GB:M77003; NID:9193366; PIDN:AAA37647.1; PID:9193367

A:Experimental source: liver

Query Match 93.1%; Score 3985.5; DB 2; Length 827;

Best Local Similarity 92.5%; Pred. No. 7.3e-295;

Matches 766; Conservative 32; Mismatches 23; Indels 1; Gaps 1;

QY	1	MDESALTGTIDVSYLPHSSSYSGRCRKHTEEMGECGFRPTVFRSATLKWKESLMSRKR	60
DB	1	MEESVTVGTIDVSYLPSSESYSLGRCKHTESEWDVDCGKPTFFRSATLKWKESLMSRKR	60
QY	61	PFVGRCCYSCYTSQSDWKFNFSISLGRNVIYINETHTRHGWLARLSYLVLFQERDV	120
DB	61	PFVGRCCYSCYTSQSDWKFNFSISLGRNVIYINETHTRHGWLARLSYLVLFQERDV	120
QY	121	HKGMPATVNTENLNSRVQEAIAEVAELNPDGSAQQSKAVNKKKAKRILQEMVAT	180
DB	121	HKGMPATVNTENLNSRVQEAIAEVAELNPDGSAQQSKAVNKKKAKRILQEMVAT	180
QY	181	VSPAMIRLTGVLLKLFNSFFNIIQHKQLEWKAATETNLLPLFLPVHRSHIDYLLLT	240
DB	181	VSPGMRITGVLLKLFNSFFNIIQHKQLEWKAATETNLLPLFLPVHRSHIDYLLLT	240
QY	241	FILFCHNFKAPYIAGNNLNIPFSTLIHKLGFFIRRLDETPDGRKDVLRALLHGHV	300
DB	241	FILFCHNFKAPYIAGNNLNIPFSTLIHKLGFFIRRLDETPDGRKDVLRALLHGHV	300
QY	301	VELLRQQQFLFLEGTGRSRGKTSRAGLLSVVVDLTSTNVIPIILLIPVIGISYDRIT	360
DB	301	VELLRQQQFLFLEGTGRSRGKTSRAGLLSVVVDLTSTNVIPIILLIPVIGISYDRIT	360
QY	361	EGHYNGEQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLESQSKPVS	420
DB	361	EGHYNGEQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLESQSKPVS	420
QY	421	LLSLEQALLPAILSPRPSDADEGDTISNSRNATDSLRRLRIANLAHILFTASKSC	480
DB	421	PLSLFQALLPAILSPRPNDVAHQDLSNESRNPDAEFRRLRIANLAHILFTASKSC	480



481 AIMS... 540  
481 AIMS... 540  
541 GNC... 600  
541 GNC... 600  
601 GPT... 660  
601 GSA... 660  
661 HD... 720  
661 QD... 719  
721 LQ... 780  
720 LQ... 779  
781 NAV... 828  
780 NAV... 827

RESULT 2  
T29448  
hypothetical protein F08F3.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T29448  
R:Blanchard, M.; Bradshaw, H.  
submitted to the EMBL Data Library, July 1996  
A:Description: The sequence of C. elegans cosmid F08F3.  
A:Reference number: 220620  
A:Accession: T29448  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-718 <BLA>  
A:Cross-references: EMBL:U64847; PIDN:AAB04876.1; GSPDB:GN00023; CESP:F08F3.2  
A:Experimental source: strain Bristol N2; clone F08F3  
C:Genetics:  
A:Gene: CESP:F08F3.2  
A:Map position: 5  
A:Introns: 42/3; 65/2; 156/3; 279/1; 310/3; 346/3; 406/1; 516/1; 552/3; 600/3; 667/1

Query Match 15.8%; Score 678; DB 2; Length 718;  
Best Local Similarity 27.2%; Pred. No. 3.2e-43;  
Matches 197; Conservative 148; Mismatches 264; Indels 114; Gaps 21;

128 NVT... 187  
78 NVE... 124  
188 LTG... 247  
125 CCS... 184  
248 IKAP... 307  
185 LKL... 244  
308 QFL... 367  
245 MPI... 304  
368 QLG... 425  
305 LMG... 359  
426 QAL... 475  
360 HRT... 406

8780/421

476 ASK... 531  
407 AQM... 466  
532 VNM... 591  
467 LVK... 519  
592 AVL... 651  
520 ---... 566  
652 QYG... 708  
567 ---... 601  
709 SOS... 763  
602 RDA... 659  
764 ERN... 817  
660 DLP... 713  
818 LLE 820  
714 VLE 716

RESULT 3  
B82732  
glycerol-3-phosphate acyltransferase XF1031 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
C:Accession: B82732  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; PMID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: B82732  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-870 <SIM>  
A:Cross-references: GB:AE003940; GB:AE003849; NID:g9105966; PIDN:AAF83841.1; GSPDB:GN0012  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.  
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohme  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigre  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaka  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; da Silveira  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF1031  
C:Superfamily: glycerol-3-phosphate acyltransferase

Query Match 11.8%; Score 505; DB 2; Length 870;  
Best Local Similarity 22.4%; Pred. No. 6.4e-30;  
Matches 197; Conservative 148; Mismatches 314; Indels 222; Gaps 26;

13 VSY... 64  
64 VOY... 123

[illegible]

RESULT 5  
AB1015  
glycerol-3-phosphate O-acyltransferase (EC 2.3.1.15) - *Salmonella enterica* subsp. enter.  
C/Species: *Salmonella enterica* subsp. *enterica* serovar Typhi  
A/Note: this species has also been called *Salmonella typhi*  
C/Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C/Accession: AB1015  
R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar  
S.; Moutle, S.; O'Gaora, F.

[illegible]

RESULT 6  
B82365  
glycerol-3-phosphate acyltransferase VC0093 [imported] - *Vibrio cholerae* (strain Ni6961)  
C:Species: *Vibrio cholerae*  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: B82365  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
Chardon, D.; Ermlaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H.  
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes: of the cholera pathogen *Vibrio cholerae*.  
A/Reference number: A82035; MUID:20406833; PMID:10952301  
A/Accession: B82365  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-811 <HEI>  
A/Cross-references: GB:AE004100; GB:AE003852; NID:g9654487; PIDN:AAF93271.1; GSPDB:GN001:  
A/Experimental source: serogroup O1; strain N1961; biotype El Tor  
C/Genetics:  
A/Gene: VC0093  
A/Map position: 1  
C/Superfamily: glycerol-3-phosphate acyltransferase

	Query Match	9.3%	Score 396;	DB 2;	Length 811;
	Best Local Similarity	23.3%	Pred. No. 1.1e-21;		
	Matches 157;	Conservative 119;	Mismatches 217;	Indels 182;	Gaps 30;
Qy	47	ATLKW--KESLMSKRFPV-----	-----GRCCYSCYFQSWDKFNFSPISLGLR	89	
Db	129	ATVLWGKPKGEGERPYLQALNGPEKALAVLASGRDCL-----	-----VRFSPV-----	174	
Qy	90	NVIYINETH-----TRRGWLARLSYVLFQERDVHKG-----	-----MEATNVTVENLNS	136	
Db	175	SMRYADTHGTGDASIAHK-----LARVARLHFGRQKLAASGPNLPQRAQLFA-----	-----RLMNS	226	
Qy	137	SRVOEALAEVAEELNPDGSAOQSKAVNKVKKAKRIIQEMVATVSPAMI-----	-----RLTGWV	192	
Db	227	PAIEKATAD-----EAKSKQIPELKEARKEAHDILDEIAADFSYLVKKGDRILGWL	277		
Qy	193	LLKLFNSFFWNIIQHKQLEVMVKAATENL-----	-----PLFLPVHRSHIDYLLTLFILFCHN	247	
Db	278	-----WN-RIYQG-LNINNAATVRRLAQDGHEIVVVPCHRSMDYLLSYLVLYHEG	326		
Qy	248	IKAPYIASGNLNT-----PIESTLIHKLGGFIERRLDETPOGRKDVLYRALLHGHIVEL	303		
Db	327	MVPPHIAAGNLNFPFAGPIF-----RRGGAPFIRRSFKGAP-----	-----LYSTIFREYLAEL	376	
Qy	304	LRQQQTEIFLEGRRSGTKSCARAGLLSVVVDLTSTNVI	PDILII	PVGISVDRIIE-G	362
Db	377	FAKGYSEYFSEGRSGTRGRLLPDKTGMAMTIIQMLRGLNRPVTILVPVYIGYEHVNEVG	436		
Qy	363	HYNGEQLGPKKNESLMSVARGVIRMLRK--	NYGCVRVDF	FAQFSLKEYLESQSQKPVSA	420
Db	437	TYAKELRGKKEKEN-----AGLVLRTLRLNFGQGVNFGPEIPLNQELNE-----	-----	484	
Qy	421	LLSLEQALLPAILSPRSDADEGRDTSINSRNATDESRLRRLIANLAEHLFTASKSC	480		
Db	485	-----TVPQWTQIDIPMGEB--SKPQWMTPTVKNLANRMMTHINDAAAAVNATLTC	531		
Qy	481	AIMSTHIVACILLVYRHROGIDSLTILVEDFFVWKEEV---	LARDFDLGF-----	SGNSEDV	532
Db	532	A-----TALLAGRQ-----ALARDNLTKQVDCYLSLRLNVVPYSATSTLPSESAEKL	578		
Qy	533	VMHAIQL-----LGNQVTVITHTSNDEFFITPTTVPVSFVFLNFSYNGVLHVFI	583		
Db	579	VEHAESLDKFVETDTWGDIIISL-----DRNQSI	LMT-----	YRNNIIHLALP	623
Qy	584	AIATCSLYAVLNKRGGLGTTSTPPNLSIQEQLVRKAASLYLLSNEGTTISLPCQTQYVC	643		
Db	624	SLIAQLLIR-----QOSVSELEKQVATVAIQYIPFLKQELFLRFEAEELNDLV	669		
Qy	644	HETVGKFIQYIGILTV	658		
Db	670	LRVVAELAROGILTV	684		

RESULT 7  
D64090  
glycerol-3-phosphate O-acyltransferase (EC 2.3.1.15) - Haemophilus influenzae (strain Rd  
D64090  
C|Species: Haemophilus influenzae  
C|Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 05-May-2000  
C|Accession: D64090  
R|Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. A:Reference number: A64000; MUID:9350630; PMID:7542800  
A:Accession: D64090  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-810 <TIGR>  
A:Cross-references: GB:U32758; GB:L42023; NID:g1573747; PIDN:AAC22406.1; PID:g1573755; T32758  
C:Superfamily: glycerol-3-phosphate acyltransferase  
C:Keywords: acyltransferase; coenzyme A; membrane protein; phospholipid biosynthesis

```
Query Match          9.2%; Score 395; DB 2; Length 810;
Best Local Similarity 22.6%; Pred. No. 1.3e-21;
Matches 166; Conservative 132; Mismatches 274; Indels 164; Gaps 28;

QY 105 LARRLSYVL---FIQERDVHKGMFATN---VTENVLNSRRVQEAIAEVAELNPDGSAQ 158
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 186 LAQKLARVAKMFAKQISATGPRLPNQRQAMFNKLLQSEAIRRAI-----EDEAKS 236

QY 159 QSKAVNKVKAKRILQEWATVSPAMIRLTGCVLLKLPNSFFWNIQIHKGQLEWKAAT 218
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 237 KNISIEKAQKAYKILDBIAADVSHSLRAVDRFLRWLWNLKLYGINSYQNSN-RVRKLAL 295

QY 219 ETNLPFLPVRSHIDYLLTFLFCHNIKAPYIASGNLNINIPFSLIHKLGFFIRR 278
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 296 EGH-EIVVPCRHSHIDYLLSVLYHQGLVPPHIAAGINLFWPIGRFWSGAFIRR 354

QY 279 RLDETPDGRKQVLRALHGHIVELLRQQQFLEIFLEGRSRGKTSCARAGLLSVVVDT 338
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 355 ---TFKGNR--LYSAIPREVLSELFHSGYSVEYFIEGGRSRTGRLAPKGTGWSMTLOA 408

QY 339 LSTNVIPDILLIPVCISYDRIIE-CHYNGEQLKPKKNESLWSVARGVIRMLRKNYGCVR 397
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 409 LOHSQTRPISIVPVVGYVGVHELVDTYAKELRGAKEKENAGLVLR-VIKLRL-NLQGF 466

QY 398 VDFAPQFSLKEYL-----ESQSKPVSALLSLEQALLPAILLPSRPSDADEGRDTS 448
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 467 VNFGEPTLSNLSQHPFDWKEQNEERK-----QWFTPA----- 501

QY 449 INESRNATDESRLRRLIANLAEHILFTASKSCAIMSTHIVACLLYRHRQIGDLSLVED 508
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 502 -----VNNISKQVMINIKAAVNSMNLVGTALLSSRQALSRQALLSEQ 545

QY 509 FFMKEEVLARDFLGFGSGEDVVM-----HAIQLLNCVTTHTSRNDEPPT 558
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 546 LSSYQQLQNVPI-----STDVVLPNVTQAMLEHVALDRIGVLI-----EKDNFGEI 594

QY 559 PSTTVPSVFELNFSYNGVLHFIMEAIIACSLYAVLNKRGGLGPTSTPPNLISQEQVLRK 618
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 595 VRLERSAVLMTYYRNNIQLHVLPSLVS---ILHYEAL-----QKDLLED 639

QY 619 AASLCY-----LLSNEGTSIPCOFTFYQVCHETVGKFTIQYGILTVAEHDQDIDSPS 670
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 640 AIRKIYPPLOGLFLHFNEDLNQV-----IHQINEFARQSVIN--SDNFLSINKS 690

QY 671 LAE--QQWDKKLPEPLWSRDEEDSDFGBEQRCYLVKS-----OSKEHQ 716
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 691 KVRILQLWSAGMLEIL-----QR--YITVTILOKQPAISRAELEKESQL 733

QY 717 FITFLQRLGLPLLEAYSSAALFVNFSGVPPEPEYLOKLHKYLITRTERNVAVAE--SA 774
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 734 VAQRLSVLHGINAPEFFDKAVF--SSFIANLKEQRYFDESQVTVLDKIBELASTLSHLIST 792

QY 775 TYCL-VKNAVWFKDI 789
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 793 EICLTVKGTIEKSED 808
```

RESULT 8  
XUECAG

glycerol-3-phosphate O-acyltransferase (EC 2.3.1.15) - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C:Date: 17-May-1995 #sequence revision 17-May-1995 #text\_change 01-Mar-2002  
C:Accession: A00565; C42956; H65211  
R:Lightner, V.A.; Bell, R.M.; Modrich, P.  
J. Biol. Chem. 258, 10856-10861, 1983  
A:Title: The DNA sequences encoding plsB and dkg loci of Escherichia coli.  
A:Reference number: A92393; MUID:83291031; PMID:6309817  
A:Accession: A00565  
A:Molecule type: DNA  
A:Residues: 1-827 <LIG>  
A:Note: this sequence was partially confirmed by protein sequencing  
R:Nichols, B.P.; Green, J.M.  
J. Bacteriol. 174, 5309-5316, 1992  
A:Title: Cloning and sequencing of Escherichia coli ubiC and purification of chorismate  
A:Reference number: A42956; MUID:92355505; PMID:1644758  
A:Accession: C42956  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 805-827 <NIC>  
A:Note: sequence extracted from NCBI backbone (NCBIN:110475, NCBI:P:110483)  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: H65211  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-827 <BLAT>  
A:Cross-references: GB:AB000477; GB:U00096; NID:g2367338; PIDN:AAC77011.1; PID:g1790474;  
A:Experimental source: strain K-12, substrain M61655  
C:Genetics:  
A:Gene: plsB  
A:Map position: 92 min  
A:Function:  
C:Description: this membrane-bound enzyme catalyzes the first step in de novo phospholipid  
cid; it may also function in the regulation of membrane biogenesis  
A:Pathway: phospholipid biosynthesis  
C:Superfamily: glycerol-3-phosphate acyltransferase  
C:Keywords: acyltransferase; coenzyme A; membrane protein; phospholipid biosynthesis

```
Query Match          9.1%; Score 391.5; DB 1; Length 827;
Best Local Similarity 24.0%; Pred. No. 2.6e-21;
Matches 147; Conservative 118; Mismatches 248; Indels 99; Gaps 21;
```

```
QY 76 DKFNPSIPSLGLRNVIYINETHTRHGWLARRLSYVLFIQERDVHKGOMFATNVENLVN 135
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 182 DSFVRFS-PSVSLRRMADEHGTDKTIAQKLA-RVARMHFARQLAAVGP-RLPARQDLFN 238

QY 136 SSRVQEAIAEVAELNPDGSAQQOSKAVN--KVKKAKRILOEMVATVSPAMIRLTGWVL 193
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 239 KLLASRAIAKAVED-----EARSKKISHEKAQONAIAMEETAANFSYEMIRLTDRI 291

QY 194 LKLFNSPFNNIQLHKGQLEWKAATETNLPFLPVRSHSHIDYLLTFLFCHNIKAPYI 253
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 292 GFTWRNLYQGLNVANA--ERVRLAHGCHELVYVPCRSHMDYLLLSVLYHQGLVPPHI 349

QY 254 ASGNLNIN-----PIPSTLIHKLGQFFIRRRRLDETPDGRKQVLYRALLHGHIVELLRQQ 309
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 350 AAGINLNFWPAGPIF---RRLGAFFIR---TFKGNK---LYSTVFREYLGELFSRGYS 399

QY 310 LEIFLEGTRSRSGKTCARAGLLSVVVDLTSTNVIPDILLIPVCISYDRIIE-CHYNGEQ 368
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 400 VEYVEGGRSRTGLDLPKGTLSMTIQAMLRGGTRPTLPIYIGYEHVNEVGTAKEL 459

QY 369 LGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLES-----OSQKPVSA 422
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 460 RGATKESLPQMLRGLSKL--RNLGQYVNFGEPMPLMTYLNQHVDPWRRESIDI----- 513

QY 423 SLEQALLPAILLPSRPSDADEGRDTSINESRNATDESRLRRLIANLAEHILFTASKCAI 482
   :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 514 ---EAVRPWLTP-----TVNNIAADLVIRINNAGAA 542
```



QY 483 MSTHIVACLLLYRHROGIDLSTLVED-----FFVWKEEVLARDFDLGFSGNSEDVVMHAIQ 538  
Db 543 NAMNLCCTALLASRQSRSTREQLTQALNCYLDLMEVNPYSTDSTVP--SASASELIDHALQ 601  
QY 539 LLGNCVTITHTSRNDEFFITPTTVPVSFELNFYNGVLHVFIEMAIACSLYAVLNKRG 598  
Db 602 M--NKFEVEKOTIGDIILPREQAV-----LMTYYRNNTIAHMLVLPSLMAAIV----- 647  
QY 599 LGGTSTPPNLIISQOLVRKAASLYLLNSNETGISLPCQTFYQVCHETVGVKFIQYGIILTV 658  
Db 648 -----TQHRHISRDVLMHEVNVLPMLKAEFLRWDRDELDPVIDALANEMQROGLITL 701  
QY 659 AEHDDQEDISP 669  
Db 702 --QDELHINP 710  
RESULT 11  
D86097  
glycerol-3-phosphate acyltransferase [imported] - Escherichia coli (strain O157:H7, sube  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
A;Accession: D86097  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: D86097  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-827 <STO>  
A;Cross-references: GB:AB005174; NID:g12518990; PIDN:AAG59240.1; GSPDB:GN00145; UMG:P.Z56  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: p18B  
C;Superfamily: glycerol-3-phosphate acyltransferase  
Query Match 9.1%; Score 387.5; DB 2; Length 827;  
Best Local Similarity 24.1%; Pred. No. 5.2e-21;  
Matches 147; Conservative 116; Mismatches 249; Indels 99; Gaps 21;  
QY 76 DKFNPSIPSLGRNVIVINETHTHRWGLARRLSYLVFIQERDVHKMFATNTVENVLN 135  
Db 182 DSFVRFS--PSVSLRRMADEHGTDKTIAQKLA-RVARMHFARQLAAVGP-RLPARQDLFN 238  
QY 136 SSRVQEAETAVAELNPDGSAQOOSKAVN--KVKKAKRILOEMVATVSPAMIRLTGHWL 193  
Db 239 KLLASRAKAVED-----EARSKTIHBAQOQNAJALMEETAAANFSYMIRLTDRI 291  
QY 194 LKLFNSFPWNIQIHKGQLEMKAAETETNLPLFLPVHRSHIDYLLTLTFLFCHNIKAPYI 253  
Db 292 GFTWNLVQGINVNA--ERVQLAHGDGHELVYVPCHRSHMDYLLLSVLVHQGLVPHI 349  
QY 254 ASGNMNI-----PIPTLIHLKGGFFIRRLDETPDGRKDVLYRALLHGHIVELLRQOQF 309  
Db 350 AAGINLNPWAGPIF---RRLGAFPIR---TFKGNK--LYSTVPREYLGELFSRGYS 399  
QY 310 LEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPIILIPVGISYDRIIE-CHYNGEQ 368  
Db 400 VEYFVEGSRSTRGLDDPKTQTSMTIOAQLRGGTRPTILPIYIGYEHVMEVGTAKEL 459  
QY 369 LGKPKKNESLWSARGVIRMLRNKNGCVRVDFAPQPSLKEYLES-----OSQKPVSA 422  
Db 460 RGATKEKESLPQMLEGLSKL--RNLGQYVNFGEPMPLMTYLNQHVDPWRESIDPI----- 513  
QY 423 SLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHEHILFTASKSCAI 482  
Db 514 --EAVRPALWLP-----TVNNIAADLMVRINNAGAA 542  
QY 483 MSTHIVACLLLYRHROGIDLSTLVED-----FFVWKEEVLARDFDLGFSGNSEDVVMHAIQ 538  
Db 543 NAMNLCCTALLASRQSRSTREQLTQALNCYLDLMEVNPYSTDSTVP--SASASELIDHALQ 601

QY 539 LLGNCVTITHTSRNDEFFITPTTVPVSFELNFYNGVLHVFIEMAIACSLYAVLNKRG 598  
Db 602 M--NKFEVEKOTIGDIILPREQAV-----LMTYYRNNTIAHMLVLPSLMAAIV----- 647  
QY 599 LGGTSTPPNLIISQOLVRKAASLYLLNSNETGISLPCQTFYQVCHETVGVKFIQYGIILTV 658  
Db 648 -----TQHRHISRDVLMHEVNVLPMLKAEFLRWDRDELDPVIDALANEMQROGLITL 701  
QY 659 AEHDDQEDISP 669  
Db 702 --QDELHINP 710  
RESULT 12  
T45238  
probable transferase [imported] - Mycobacterium leprae  
C;Species: Mycobacterium leprae  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000  
A;Accession: T45238  
R;James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, March 1999  
A;Reference number: 222949  
A;Accession: T45238  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-775 <JAM>  
A;Cross-references: EMBL:AL049913; PIDN:CAB43153.1  
A;Experimental source: cosmid B1610  
C;Genetics:  
A;Note: p18B

Query Match 6.8%; Score 289; DB 2; Length 775;  
Best Local Similarity 23.6%; Pred. No. 1.5e-13;  
Matches 102; Conservative 86; Mismatches 184; Indels 60; Gaps 10;  
QY 168 KKAKRILOEMVATVSPAMIRLTGHWLLKLFNSFPWNIQIHKGQLEMKAAETETNLPLFL 227  
Db 207 EEACKMLDELATGWSRASVDLVSLGRMLSRGFPEIDYDEYQVAAMEALEAH-PAVLL 265  
QY 228 PVHRSHIDYLLTLTFLFCHNIKAPYIASGNLNIPIPTLIHLKGGFFIRRLDETPDGR 287  
Db 266 FSHRSYIDGAVVPVAMQENRLPPVHVFAGINLSFGLMGLRLRRSGVPIFERNIGDNP--- 322  
QY 288 KDVLRYALLHGHIVELLRQOQFLEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPI 347  
Db 323 ---LYKVLREYGYIVKEKPNLSWSEGTNRSTRTGKMLPPKGLGLTYVADAYLDGRSEDI 379  
QY 348 LIIPVGISYDRIIE-CHYNGEQLGKPKKNESLWSARGVIRMLRNKNGCVRVDFAPQPSL 406  
Db 380 LLQPVISIFDQLHETAETAAVARGGEKTPEGVAMLYSFIKAQGERNYGKIYVRPEAVSM 439  
QY 407 KEYLESQKPVSAALLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIA 466  
Db 440 RQYLGA---PHGALVQ-----DQDAKRLAQ 462  
QY 467 NLAEHILFTASKSCAINTHIVACLLLYRHROGIDL-----TLVE--DFFWKEEVLARD 520  
Db 463 KMSFEVNRILCAPVTATALLVSALLTTTRGVALTLQOLHHTLQESLDYLERKOTPVSKS 522  
QY 521 FDLGFSGNSEDVVMHAIQQL--GNCVTITHTSRNDEFFITPTTVPVSFELNFYNGVLH 578  
Db 523 ---ALRLSRREGVRAVDALSSGHPIITRVDSGREPVWYITPGNEHAAA-----FYRNSVIH 575  
QY 579 VFMEAIACSL 590  
Db 576 AFLETSIVELAL 587

RESULT 13  
A70868  
probable p18B2 - Mycobacterium tuberculosis (strain H37RV)  
C;Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C/Accession: A70868  
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A/Reference number: A70500; MUID:98295987; PMID:9634230  
C/Accession: A70868  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-789 <COL>  
A/Cross-references: GB:AL021246; GB:AL123456; NID:g3261507; PIDN:CAA16059.1; PID:e123760  
A/Experimental source: strain H37Rv  
C/Genetics:  
A/Gene: plsB2

Query Match 5.7%; Score 245; DB 2; Length 789;  
Best Local Similarity 22.0%; Pred. No. 3.4e-10;  
Matches 112; Conservative 95; Mismatches 206; Indels 96; Gaps 16;  
QY 116 QERD---HKGMPATNTVNLSRVQEAIAEVAE-----LMPDGSAAQ 158  
DB 149 QWRD---HKGMPATNTVNLSRVQEAIAEVAE-----LMPDGSAAQ 158  
QY 159 QSKA-VNKKV---KKAIRILOENV---ATVSPAMIRLTGWLKLFNSFFWNIQHKGO 210  
DB 201 RFRAGLDRIICATVEDAGKMLDELSTGWSQVSDVSLGRLASRGDPDEF---DYDEYQ 257  
QY 211 LEMVKAATNTNLPFLFPHRSHIDYLLTFILFCHNIKAPYIASGNLNIPFSTLIHK 270  
DB 258 VAAARAALAH-PAVLLFHSRYSIDGVVPMQDNRLPPVHFGGINSFGLMGPLMR 316  
QY 271 LGGFFTRRRDDEPDGRKOVLYRALLHGHIVELLROQQFLFIELEGRTRSRSGKTSARAG 330  
DB 317 SGMIFFTRRNNDP-----LYKVLKEYVGVYVKKFNLWSWTEGTRSRTRGMLPPKLG 370  
QY 331 LLSVVVDLTNTNIPDILLIPVGSYDRIIE-GHYNQEQLGKPKKNSLSVARGVIRML 389  
DB 371 LMSVADAYLDGRDDILLQGVSLFCFQDLHEITEYAYARGAETPRLGLRWLFNFKAQG 430  
QY 390 RKNYGVVRDFAQFSLKYLESOQKPVSAALLSLEQALLPAILPSRPSDADEGRDTSI 449  
DB 431 ERNFGKIYVAFPEAVSNQYL-----GAPHELQ- 460  
QY 450 NESRNATDESRLRLANLAHLFTASKSCAIMSTHIVACLLYRHRQIDISTLVE-- 507  
DB 461 -----DPAAKRLALQKMSFEVAWRLQATPVTATGLVSALLITTRGTALTLDLHHTL 513  
QY 508 -----DPFVMEVLRDPLDGFSGNSDVMHAIQLL--GNCVTIHTSRNDEFFITPST 561  
DB 514 QDSLDYLERKQSPVSTS-----ALRLRSREGVRAAADALSNHGPVTRVDSGREPVWYIAPDD 570  
QY 562 TVPSVFELNYSNGVLHVFIIMEAIIACSL 590  
DB 571 EHAAA-----PYRNSVIAHLETSIVELAL 595

RESULT 14  
D70762  
probable transferase - Mycobacterium tuberculosis (strain H37Rv)  
C/Species: Mycobacterium tuberculosis  
C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C/Accession: D70762  
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A/Reference number: A70500; MUID:98295987; PMID:9634230  
C/Accession: D70762

A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-621 <COL>  
A/Cross-references: GB:Z74020; GB:AL123456; NID:g3261584; PIDN:CAA98312.1; PID:e248544; I  
A/Experimental source: strain H37Rv  
C/Genetics:  
A/Gene: plsB1

Query Match 5.4%; Score 229; DB 2; Length 621;  
Best Local Similarity 19.4%; Pred. No. 3.8e-09;  
Matches 134; Conservative 116; Mismatches 270; Indels 170; Gaps 24;  
QY 126 ATNVTG---NVLSNRRVQEAIAEVAELNPDGSAQQSKAVNKYKKAIRILOENVATVSP 183  
DB 29 ATDPVEVTQLDARWYDERLRALALGRD-----PDSVRAEAAGYREMAASLDE 79  
QY 184 AMIR-----ITGVVLLKLFNSFFWNIQHKGOLEMWKAAATETNLPILFLPVRHSHIDYLL 239  
DB 80 RAVQAWRGFSRWLMRA-----YDVLVDEDOITQLR-KLDRKATLAFASFHRSYLDGMLL 132  
QY 240 TFIPLFCHNIKAPYIASGNLNIPFSTLIHKLGFFIRRRRLDTPDGRKOVLYRALLHGH 299  
DB 133 PEAILANRLSPALTFGCANLNFFPMGAWAKRTGAIFIRQTKDIP-----VTRFVLRAY 186  
QY 300 IVELLRQOQFLFIELEGRTRSRSGKTSARAGLLSVVVVDLTNTNVPIDILLIPVGSYDRI 359  
DB 187 AAQLVQNHVNLTWISIEGRTGKLRPPVFGILRYITDAVDEIDGPEVVLVPTSIVVDQL 246  
QY 360 IEGH-VNGQLGPKPKKNSLSVARGVIRMLR---KNYGVVRDFAQFSLKYLESOQS 415  
DB 247 HEVEAMTTAYGAVKRPEDL---REFVRLARQOGERLGRAYLDGPEPLRKLQ--- 298  
QY 416 KPVSAALLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLANLAHLFT 475  
DB 299 -----EMRADKSGTSGEIER--IALDVEHRI-- 322  
QY 476 ASKSCAIMSTHIVACLLYRHRQIDISTLVEFVFMKEEVLARDPLDGFSGNSDVMH 535  
DB 323 -NRATPVTPTAVVSLALIGADR-SLSISEVLATVRPLASVIAARNWAVAGAAD----- 373  
QY 536 AIQLGNCVTIHTSRNDEFFITPSTVSVFELN-----FVSNGLVHV 579  
DB 374 ----LTRNTRIRTWLHQ-----MVASGVSVVDAGTAVWVGIGEDQHLVAAAFVNTAIHI 424  
QY 580 FIMEAIIACSLYAVLNKRLGSGPTSPNLSIQEQLVRKAASCLYLLSNEGTTISLPCQTF 639  
DB 425 LVDRVAELALLAAATTTNG---SVSPATVRDE-----ALSRLDLKKEFLFSGRAQFE 476  
QY 640 YQVCHET--VGKFIQYGIILTVAEHDDQEDISPSLABEQQWKKLPEFLSMRSDDEEDSDF 697  
DB 477 KOLANEVLLIGSVV-----DTSKPAADAVM----- 502  
QY 698 GEEQRCYLVKSQKHEHQFIFLQRLGLPLEAYSAA--IFVHNPFGVPPEYLOKL 755  
DB 503 -----RLLESAD-----VLLAHVLVLPFLDAYHIVADRLAAHEDDSDPDEEGFLAECL 549  
QY 756 HKYLITRERNVAVYAESATYCLVKNVKNM 785  
DB 550 QVGKQWELQORNTIA-SAESRSMELFKTALRL 578

RESULT 15  
AG2414  
hypothetical protein all4871 [imported] - Nostoc sp. (strain PCC 7120)  
C/Species: Nostoc sp. PCC 7120  
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C/Accession: AG2414  
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S. DNA Res. 8, 205-213, 2001  
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A/Reference number: AB1807; MUID:21595285; PMID:11759940

Search completed: January 10, 2004, 20:21:45  
Job time : 40 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2004, 20:17:17 ; Search time 27 Seconds  
(without alignments)  
1297.533 Million cell updates/sec

Title: US-09-935-290-2

Perfect score: 4280

Sequence: 1 MDESALTGLTIDVSYLPHSS.....FLPCNQKQLLEYILSFVVL 828

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	462	10.8	865	US-09-252-991A-20894	Sequence 20894, A
2	438.5	10.2	862	US-09-328-352-6286	Sequence 6286, Ap
3	140	3.3	3144	US-08-246-982A-6	Sequence 6, Appli
4	140	3.3	3144	US-08-453-265-6	Sequence 6, Appli
5	140	3.3	3144	US-08-457-273B-42	Sequence 42, Appli
6	140	3.3	3144	US-08-556-419-21	Sequence 21, Appli
7	140	3.3	3144	US-09-041-886-15	Sequence 15, Appli
8	115.5	2.7	3118	US-08-457-273B-8	Sequence 8, Appli
9	115.5	2.7	3119	US-08-246-982A-16	Sequence 16, Appli
10	115.5	2.7	3119	US-08-453-265-16	Sequence 16, Appli
11	109.5	2.6	1698	US-09-315-793-12	Sequence 12, Appli
12	105.5	2.5	2188	US-09-328-352-7763	Sequence 7763, Ap
13	103	2.4	1356	US-09-770-170-6	Sequence 6, Appli
14	102	2.4	531	US-09-252-991A-24404	Sequence 24404, A
15	100.5	2.3	2930	US-09-417-822-2	Sequence 2, Appli
16	100	2.3	2108	US-09-252-991A-31502	Sequence 31502, A
17	99.5	2.3	3665	US-08-222-617A-13	Sequence 13, Appli
18	99.5	2.3	3712	US-08-222-617A-25	Sequence 25, Appli
19	98.5	2.3	383	US-08-484-575A-4	Sequence 4, Appli
20	98.5	2.3	383	US-08-477-459-4	Sequence 4, Appli
21	98.5	2.3	383	US-08-478-869-4	Sequence 4, Appli
22	98.5	2.3	383	US-08-486-414-4	Sequence 4, Appli
23	98.5	2.3	383	PCT-US94-01826A-4	Sequence 4, Appli
24	98.5	2.3	383	PCT-US94-02252A-4	Sequence 4, Appli
25	98.5	2.3	840	US-08-974-549A-190	Sequence 190, App
26	98.5	2.3	872	US-08-851-843A-8	Sequence 8, Appli
27	98.5	2.3	872	US-08-851-843A-54	Sequence 54, Appli

28	98.5	2.3	872	3	US-08-974-549A-221	Sequence 221, App
29	98.5	2.3	872	3	US-08-854-050-8	Sequence 8, Appli
30	98.5	2.3	872	3	US-08-854-050-54	Sequence 54, Appli
31	98.5	2.3	872	4	US-09-430-323-8	Sequence 8, Appli
32	98.5	2.3	872	4	US-09-430-323-54	Sequence 54, Appli
33	98.5	2.3	3712	2	US-08-222-617A-4	Sequence 4, Appli
34	98	2.3	1745	2	US-09-031-485-33	Sequence 33, Appli
35	98	2.3	1745	2	US-08-847-429A-33	Sequence 33, Appli
36	98	2.3	1745	3	US-09-065-474-33	Sequence 33, Appli
37	98	2.3	1745	4	US-09-557-034-33	Sequence 33, Appli
38	97.5	2.3	303	2	US-08-321-670-2	Sequence 2, Appli
39	97.5	2.3	303	4	US-09-215-252-3	Sequence 3, Appli
40	96.5	2.3	942	2	US-08-884-681-4	Sequence 4, Appli
41	96.5	2.3	942	3	US-09-258-643-4	Sequence 4, Appli
42	96	2.2	317	4	US-09-107-532A-3993	Sequence 3993, Ap
43	95.5	2.2	395	3	US-09-080-044-5	Sequence 5, Appli
44	95.5	2.2	395	4	US-09-531-857A-5	Sequence 5, Appli
45	95	2.2	564	2	US-08-948-569A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1  
US-09-252-991A-20894  
; Sequence 20894, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20894  
; LENGTH: 865  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20894

Query Match	10.8%;	Score 462;	DB 4;	Length 865;
Best Local Similarity	22.8%;	Pred. No. 6.3e-39;		
Matches 171;	Conservative 150;	Mismatches 262;	Indels 166;	Gaps 27;
QY	88	LRNVIYINETHTRHGRGLARRLSYVLFQERDVHKGHFA-----TNVTENVLNSSRYOE	141	
DB	209	LRELVEQKGHER-----TLRMVNRILRVHFNLTAVIGPDLSHRNLVKGLLRAPLVYRQ	264	
QY	142	AIAEVAALNPDGSAQOOSKAVNKVKKAKRILQEMVATVSPAMIRLTGWVLLKLFNSFF	201	
DB	265	AISE---ECESERISQEKAGI-----ALRYANEAISDFSPVIR-----FLEVILSWF	310	
QY	202	WNIQIHKG-----QLEMYKAATETNLPFLFVPHRSHDYLILLTFILFCHNFKAPIYASGN	257	
DB	311	WN-KLYEGVKVNHIERVQDVAQGN-EIVYVPCHRSHDYLILLTFILFCHNFKAPIYASGN	368	
QY	258	NLNPIPISTLHKLGFFIRRRRLDTPDGRKDVLYRALLHGHIVELLRQQQFLFLEFGT	317	
DB	369	NLNPVIGSLIRRGGAFFMRSP-----KGNQLYTAVNFVNLTLFSGRSTFVVEGG	422	
QY	318	RSRSGKTSARAGLLSVVVDTLSTNVIPIIDILIIIPVGISYDRIIEGH-YNGBQLGPKKNE	376	
DB	423	RSRTGRMLHPRTGLAITLRSFLDRSRRIIVFVFPVYGYERVLGRTYLGELRGATKKCE	482	
QY	377	SLWSVAGVTRMLRKNYGCVRVDFPAQPSLKEYLESQSKPVSAALLSLEQALLPAILPS-	435	
DB	483	SIFDLFK-VVGALKQRFQGVWVNFGEPTHLDDQFLDRHQ-----PDW	522	



INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3144 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-246-982A-6

Query Match 3.3%; Score 140; DB 1; Length 3144;

Best Local Similarity 18.7%; Pred. No. 0.00048;

Matches 160; Conservative 130; Mismatches 319; Indels 248; Gaps 40;

```
QY 57 SRKPPFVGRCCYCTPQSWDKFPNPSIPSLGRNVYINETHRHGWLARRLSYVLFQ 116
DB 1350 SSVRPGLYHYCF-MAPYT---HFTQALADASLRNMVQA-EQENDTSGWFD-----VLQ 1397

QY 117 ERDVHKGMFATNTVEN-----VLNRRVQEAIAEVAELNPDGSAQQSKAVNKKVKA 170
DB 1398 KVTQLKNTLTSVTKRADKNAIHNHILRPELVIKALK-----QYTTTCVOLQKQV 1450

QY 171 KRILQEMVATVSPAMIRLTGWL-----LKLNSFFWNIIQHKQLEMVKAATET 220
DB 1451 LDLLAQLV-----QLRVNYCLLSDQVFIGFVLKQFE-----YIEVGQFR---ESEA 1494

QY 221 NLP-----LLFLPVRSHIDYL--LLTFILFCHNIKAPYIASGNL---NIPFSTLIHK 270
DB 1495 IIPNIFFLVLSYVYHSHKQIIGIPKIIQLCDGI---MASGRKAVTHAIPALQPIVHD 1550

QY 271 LGGFFIRRLDETDPGRKDVLYRALLHGHIVELLRQQQFLFIELEGTRSRSGKTSARAG 330
DB 1551 L---FVLGRTNKADAGKELETQKEVVVSMRLRLIYHQVLEMF-----KENEKWK-----RLSR 1610

QY 331 LLSVVVDLTSTNVPIDILIIIPVGISYDRIIEGHYNGEQLGPKPKNESLWSVARGVIRMLR 390
DB 1591 -----ILVLQ-----QCH-----KENEKWK-----RLSR 1610

QY 391 KNYGCVRVDFAPQPSLKEYLESQKPVSALESQALLPAILPSRPSDADEGRDTSIN 450
DB 1611 Q-----IADIILPMLAKQMHDSHEALGVLTILFEILAPSSL--RPVDMLLRSMFVTPN 1663

QY 451 ESRNATDES-----RRRLIANLAELHILFASKSCAINTHIVACLLIYHRQGDIL 502
DB 1664 TMASVSTVQLMISGILAIRLVLIQSOTEDIVLSRIQELS-FSPYLISCTVINRLRDGST 1722

QY 503 STLVE-----DFFVMKEEVLARDFDLGFG-NSEDEVNHAITQLLGNCTVITHTSRNDEFF 556
DB 1723 STLEHSGKQIKNLPEETFSR-FLQLVGLLEDIVTKLK-----VEMSEQOHTFY 1774

QY 557 ITPSTVPSVELNPFYSGNVLHVFI--MEAIACSLYAVLNKRGLGGTSTPPNLSIQEQ 614
DB 1775 CQ-----ELGTLMLCLIHIFKSGMFRITAAATRLFRSDGCG-----SPYTLDS 1819

QY 615 LVRAASLCYLLSNEGTSISLPCQFYQVCHETVKGFIQYGLITVAEHDDQEDISPSLAEQ 674
DB 1820 LNLKARS--ITTHPALVILWCQ-----ILLVNHDTY-----R 1851

QY 675 QWDKKLPEP-----LSWRSDDEDESDFGEEQDCYLK-----VSQS 711
DB 1852 WVAEVQQTPKRHSLSSTLLSPQSGEEDSDLAALGMCNREIVRGCALILFCDYVCQN 1911

QY 712 KEHQOFTFL-----QRLLG-----PLLEAYSSA-----AIFV-----HNFSGVPV 747
DB 1912 LHDSEHLTWLVNHIQDLISLSHEPPVDFTISAVHRNSAASGLFIQAIQSCENLSPTM 1971

QY 748 EPEVLQKHLKYLITERNVAVYAESATYCLVKAVKRMFKDIGVFKETQKRVSVLESS 807
DB 1972 LKKTQLQCEGHLQSOGAVLTYVD-RLLCPTFRVLARMVDILACRRVEMLLAANLQSSM 2030

QY 808 TFLPQCNRQKLEYILS 824
DB 2031 AQLPWEELNRQIEYLOS 2047
```

#### RESULT 4

US-08-453-265-6

; Sequence 6, Application US/08453265

; Patent No. 5693757

; GENERAL INFORMATION:

; APPLICANT: MacDonald, Marcy E.

; APPLICANT: Ambrose, Christine M.

; APPLICANT: Duyao, Mabel P.

; APPLICANT: Gusella, James F.

; TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox

; STREET: 1100 New York Avenue

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/08/453,265

; FILING DATE: 30-MAY-1995

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Ludwig, Steven R.

; REGISTRATION NUMBER: 36,203

; REFERENCE/DOCKET NUMBER: 0609.3880003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3144 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-453-265-6

Query Match

3.3%; Score 140; DB 1; Length 3144;

Best Local Similarity 18.7%; Pred. No. 0.00048;

Matches 160; Conservative 130; Mismatches 319; Indels 248; Gaps 40;

```
QY 57 SRKPPFVGRCCYCTPQSWDKFPNPSIPSLGRNVYINETHRHGWLARRLSYVLFQ 116
DB 1350 SSVRPGLYHYCF-MAPYT---HFTQALADASLRNMVQA-EQENDTSGWFD-----VLQ 1397

QY 117 ERDVHKGMFATNTVEN-----VLNRRVQEAIAEVAELNPDGSAQQSKAVNKKVKA 170
DB 1398 KVTQLKNTLTSVTKRADKNAIHNHILRPELVIKALK-----QYTTTCVOLQKQV 1450

QY 171 KRILQEMVATVSPAMIRLTGWL-----LKLNSFFWNIIQHKQLEMVKAATET 220
DB 1451 LDLLAQLV-----QLRVNYCLLSDQVFIGFVLKQFE-----YIEVGQFR---ESEA 1494

QY 221 NLP-----LLFLPVRSHIDYL--LLTFILFCHNIKAPYIASGNL---NIPFSTLIHK 270
DB 1495 IIPNIFFLVLSYVYHSHKQIIGIPKIIQLCDGI---MASGRKAVTHAIPALQPIVHD 1550

QY 271 LGGFFIRRLDETDPGRKDVLYRALLHGHIVELLRQQQFLFIELEGTRSRSGKTSARAG 330
DB 1551 L---FVLGRTNKADAGKELETQKEVVVSMRLRLIYHQVLEMF-----KENEKWK-----RLSR 1610

QY 331 LLSVVVDLTSTNVPIDILIIIPVGISYDRIIEGHYNGEQLGPKPKNESLWSVARGVIRMLR 390
DB 1591 -----ILVLQ-----QCH-----KENEKWK-----RLSR 1610

QY 391 KNYGCVRVDFAPQPSLKEYLESQKPVSALESQALLPAILPSRPSDADEGRDTSIN 450
DB 1611 Q-----IADIILPMLAKQMHDSHEALGVLTILFEILAPSSL--RPVDMLLRSMFVTPN 1663
```



GENERAL INFORMATION:  
; APPLICANT: Ross, Christopher  
; APPLICANT: Li, Xiao-Jiang  
; APPLICANT: Li, Shi-Hua  
; APPLICANT: Sharp, Alan  
; APPLICANT: Lanahan, Anthony  
; APPLICANT: Worley, Paul  
; APPLICANT: Snyder, Solomon  
; TITLE OF INVENTION: Huntingtin-associated protein  
; FILE REFERENCE: 01107.52271  
; CURRENT APPLICATION NUMBER: US/08/556,419C  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PsetSeq for Windows Version 3.0  
; SEQ ID NO 21  
; LENGTH: 3144  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-556-419-21

Query Match 3.3%; Score 140; DB 3; Length 3144;  
Best Local Similarity 18.7%; Pred. No. 0.00048;  
Matches 160; Conservative 130; Mismatches 319; Indels 248; Gaps 40;

QY 57 SRKRPFGRCYCCTPQSDWDFNPSIGLNVVINYETHRHGWLARRLSVLFPIQ 116  
DB 1350 SSVRPGLYHYCF-MAPYT---HFTQALADASLRNMVQA-EQENDTSGWFD-----VLQ 1397

QY 117 ERDVHKGMFATNVTEEN-----VLNSSRVQEAIAEVAELNPDGSAQOQSKAVNVKKA 170  
DB 1398 KVSQKLNLSVTNKRADKNAHNRHRLPELVKALK-----QYTTTCVQLQKQV 1450

QY 171 KRILQEMVATVSPAMIRLTGWL-----LKLNSFFWNIIQHKQLEVMVKAATET 220  
DB 1451 LDLLAQLV-----QLRVNYCLDSDQVFGVLKQFE-----YIEVGQFR-----ESEA 1494

QY 221 NLP-----LLFLPVHRSHIDL--LLTFILFCHNIKAPYIASGNL---NIPFSTLIHK 270  
DB 1495 IIPNIFFLVLLSYRHSKQIIGIPKIIQLCDGI---MASGRKAVTHAIPALQPIVHD 1550

QY 271 LGGFPIRLDETDPGRDVLRYALLHGHVLLRQQQFLIFLEIGTRSRSKGTSCARAG 330  
DB 1551 L---FVLRTKADAGKELETKQEVVSMMLRLIYHQVLEMF-----1590

QY 331 LLSVVVDTLSTNVIPDIIIPGISYDRIIEGHVNGEQLGPKPKNESLWSVARGVIRMLR 390  
DB 1591 -----ILVLQ-----QCH-----KENEDKWK-----RLSR 1610

QY 391 KNYGCVRVDFAPFSLKEYLESQKPVSALESLEQALLPAILPSPSDAADEGRDTSIN 450  
DB 1611 Q-----IADIILPMLAKQMHDSHEALGVNLTLPFILAPSSL--RPVDMLLRSMFVTPN 1663

QY 451 ESRNATDESL-----RRLIANLAHILFTASKSCAIMSTHIVACLLYRHRQGLD 502  
DB 1664 TMASVSTVOLMISGILAILRVLISOSTEDIVLSRIQELS-FSPVLSICTVNRRLRDGDST 1722

QY 503 STLVE-----DFFVMKEELVARDPDLFGSG-NSEDVVMAHAIQLLNCVTTIHTSRNDEFF 556  
DB 1723 STLSEHSGKQIKNLPETFSR-FLQLVGLLEDIVTKQK-----VEMSQOHTFY 1774

QY 557 ITPSTTVSPVELNPSYNGVLHVTI--MEAIACSLYAVANKRGIGGTPSTPPNLSIQEQ 614  
DB 1775 CQ-----ELGTLMLCLIHIFKSGMFRITAAATRLFRSDGCGG-----SFYTLD 1819

QY 615 LVRKAASLYLLSNETGISLTCQFYQVCHETVGKFIQYIGILTVAEHDDQEDISPSLAEQ 674  
DB 1820 LNLARSM--ITTHPALVLLWCQ-----ILLVNHDTY-----R 1851

QY 675 QWDKKLPEP-----LSWRSDDEDDSDFGGEQDCYLK-----VSQS 711  
DB 1852 WAAEQVQTPKHSLSSTKLLSPQSGEEDSDLAAGKNCNREIVRGAALLFCDYVCQN 1911

QY 712 KEHQOQITFL-----QRLLG-----PLLEAYSSA-----AIFV-----HNFSGPVP 747

DB 1912 LHDSEHLTWLIVNHIQDLISLSHEPPVQDFISAVHRNSAASGLFIQATQSCENLSTPTM 1971

QY 748 EPEYLOKLHKYLIITRERNVAVYAESATYCLVKNVAKMFKDIGVFKETQKRVSVLELSS 807

DB 1972 LKKTLOCLEGHLSSQSGAVTLIYVD-RLCTPFRVLARMVDILACRRVEMLLAANLQSSM 2030

QY 808 TFLPQCNROKLELYLS 824

DB 2031 AQLPWEELNRIOEYLS 2047

RESULT 7  
US-09-041-886-15  
; Sequence 15, Application US/09041886  
; Patent No. 6235872  
; GENERAL INFORMATION:  
; APPLICANT: Bredesen, Dale E.  
; APPLICANT: Rabizadeh, Sharroz  
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence  
; TITLE OF INVENTION: Polypeptides and Methods of Use  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/09/041,886  
; APPLICATION NUMBER: US/09/041,886  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 2626  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3144 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-041-886-15

Query Match 3.3%; Score 140; DB 3; Length 3144;  
Best Local Similarity 18.7%; Pred. No. 0.00048;  
Matches 160; Conservative 130; Mismatches 319; Indels 248; Gaps 40;

QY 57 SRKRPFGRCYCCTPQSDWDFNPSIGLNVVINYETHRHGWLARRLSVLFPIQ 116  
DB 1350 SSVRPGLYHYCF-MAPYT---HFTQALADASLRNMVQA-EQENDTSGWFD-----VLQ 1397

QY 117 ERDVHKGMFATNVTEEN-----VLNSSRVQEAIAEVAELNPDGSAQOQSKAVNVKKA 170  
DB 1398 KVSQKLNLSVTNKRADKNAHNRHRLPELVKALK-----QYTTTCVQLQKQV 1450

QY 171 KRILQEMVATVSPAMIRLTGWL-----LKLNSFFWNIIQHKQLEVMVKAATET 220  
DB 1451 LDLLAQLV-----QLRVNYCLDSDQVFGVLKQFE-----YIEVGQFR-----ESEA 1494

QY 221 NLP-----LLFLPVHRSHIDL--LLTFILFCHNIKAPYIASGNL---NIPFSTLIHK 270  
DB 1495 IIPNIFFLVLLSYRHSKQIIGIPKIIQLCDGI---MASGRKAVTHAIPALQPIVHD 1550

QY 271 LGGFFIRRLDETPDGRKDVLYRALLGHVIELLROQQFLIEIFLEGTRSRSGTKSCARAG 330  
Db 1551 L----FVLRTGNKADAGELETKQEVVSMRLRLIYHQVLEMF-----1590  
QY 331 LLSVVDTLTNNVPIILLIIPVGISYDRIIEGHYNGEQLGPKPKNSLSWVARGVTRMLR 390  
Db 1591 -----ILVLQ-----QCH-----KENEDKWK-----RLSR 1610  
QY 391 KNYGCVRVDFAPFSLKYLESSQKPVSALESQALLPAILPSRPSDAADSGRTSIN 450  
Db 1611 Q-----IADILPMLAKQOHIDSHGALGVNLTLEILAPSSL--RPVDMLLRSMFVTPTN 1663  
QY 451 ESRNATDESL-----RRRLIANLAHILFTASKSCAIMSTHIVACILLYHRHQIDL 502  
Db 1664 TMASVSTVQLWISGILLAILRVLSQSTEDIVLSRIQELS-FSPYLISCTVINLRDGDST 1722  
QY 503 STLVE-----DFFVMKEEVLARPDGLGSG-NSEDEVVMHAIQLLGNCTVITHTSRNDEFF 556  
Db 1723 STLEHSEGGKIKNLPETFSR-FLQLVGILLEIVTKLK-----VEMSEQQHTFY 1774  
QY 557 ITPTSTTVPSVPELNFYSNGVLHVFI--MEALIASLVAVLNKRGGLGPTSTTPPNLISQEQ 614  
Db 1775 CO-----ELGTLMLCUIHIFKSGMFRITAAATRLPRSDGCG-----SFTLDS 1819  
QY 615 LVRAKASLCYLLSNEGTSISLPCQTFYQVCHETVGKFIQYGLTVAEHDDQEDISPSLAEQ 674  
Db 1820 LNLRRSM--ITTHPALVLLWCQ-----ILLVNHDTY-----R 1851  
QY 675 QWKKUPEP-----LSWRSDEDEDSFGEBQORDCYLK-----VSQS 711  
Db 1852 WMAEVOQTPKRHSLSSTKLLSPQMSGEEEDSLAAKGLMCMREIVRRGALLILFCYVCON 1911  
QY 712 KEHQQTTFI-----QLLGG-----PULLEYSSA-----AIFV-----HNFSGPVP 747  
Db 1912 LHSEHLTWLVNHIQDLSLSHEPPVQDFISAVHRNSAASGLFIOAIRSCENLSTPTM 1971  
QY 748 EPEYLQKLKYLITRTRNVAVVAESATYCLVKNVAKMFKDIGVFETKOKRVSVLELS 807  
Db 1972 LKKTQLCLEGHIHSQSAVLTLYVD-RLCTPFRVLARMVDILLACRVEMLLANLQSSM 2030  
QY 808 TFLPQCNRQKLLEYLS 824  
Db 2031 AQLPMBELNRIQEVLOS 2047

RESULT 8

US-08-457-273B-8  
; Sequence 8, Application US/08457273B  
; Patent No. 5849995  
; GENERAL INFORMATION:  
; APPLICANT: Hayden, Michael  
; APPLICANT: Lin, Biaoyang  
; APPLICANT: Nasir, Jamal  
; TITLE OF INVENTION: Mouse Model for Huntington's Disease and  
; TITLE OF INVENTION: Related DNA Sequences  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Virginia Bennett  
; STREET: PO Box 37428  
; CITY: Raleigh  
; STATE: No. 5849995th Carolina  
; COUNTRY: US  
; ZIP: 27627  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/08/457,273B  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:

NAME: Bennett, Virginia C.  
REGISTRATION NUMBER: 37,092  
REFERENCE/DOCKET NUMBER: 3477-85A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-854-1400  
TELEFAX: 919-854-1401  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3118 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-457-273B-8

Query Match 2.7%; Score 115.5; DB 2; Length 3118;  
Best Local Similarity 18.1%; Pred. No. 0.18;  
Matches 155; Conservative 128; Mismatches 321; Indels 253; Gaps 41;

QY 57 SRKRPFVGRCCYCTPQSWDKFPNPSISLGRNVIYINETHTRHGRWLARRLSYVLFQ 116  
Db 1327 SSVRGLHYCF-MAPYT---HFTQALADASLRNVQA-EQERDASGWF-----VLQ 1374  
QY 117 ERDVHKGMPATNVNEN-----VLNRRVQEAIAEVAELNPDGSAQQOSKAVNKVKKKA 170  
Db 1375 KVSQALKTTLSVTKNRADKNAIHNRILFEPLV-----IKALKQVTTTT 1419  
QY 171 KRILQEMVATVSPMIRL-TGWVLL---KLNSF-----FWNIQIHKGLEWVKAATETNL 222  
Db 1420 SVQLQKQVLDLQAOLVQLRVNYCLLSDQVFGFVLKQFEYIEV--GQFR-----ESEAI 1473  
QY 223 P-----LLFLPVHRSHIDYL--LLTFILFCHNIKAPYIASGNL--NIPIFSTLIHKL 272  
Db 1474 PNIFFVLVLSYRTHSKQIIGIKIQLCDGI-----MASGRKAVTHAIPALQPIVHDL- 1528  
QY 273 GFFIRRLDETDPGRKDVLYRALLGHVIELLROQQFLIEIFLEGTRSRSGTKSCARAGL 332  
Db 1529 --FVLRTGNKADAGELETKQEVVSMRLRLIYHQVLEMF-----1567  
QY 333 SVVDTLTNNVPIILLIIPVGISYDRIIEGHYNGEQLGPKPKNSLSWVARGVTRMLRKN 392  
Db 1568 -----ILVLQ-----QCH-----KENEDKWK-----RLSRO- 1588  
QY 393 YGCVRVDFAPFSLKYLESSQKPVSALESQALLPAILPSRPSDAADSGRTSIN 452  
Db 1589 -----IADILPMLAKQOHIDSHGALGVNLTLEILAPSSL--RPVDM--LRSMFITS 1640  
QY 453 RNATDESL-----RRRLIANLAHILFTASKSCAI-----MSTHIVACILLYHRHQ 498  
Db 1641 TMASVSTVQLWISGILLAILRVLSQSTEDIVL-----CRIQELSPHLLSCPVINLRG 1695  
QY 499 G---IDL-----STLVEDFF-----VMKEEVLARDFDLGFSNSEDVVMHAIQ 538  
Db 1696 GGGNVTLGCSRGKQKSLPEDTFSRLLQLVGLLEDIVTKQLKVDMSQQHTFYCQELG 1755  
QY 539 LLGNCVTIHTSRNDEF-FITPSTT---VPSVFELNFYSNGVLHVFI-----AII 586  
Db 1756 TLLMC--LIHIFKSGMFRITAAATRLFTSDCEBSFYTLESINARVSRMVTWTHPALVLL 1813  
QY 587 ACSLYAVLN-----KRLGGPTSTTPPNLISQBLVRKAASL--CYL-LSN 628  
Db 1814 WCQILLINHTDYRWMAEVQQTPKRHSLSCTKSLNPKSGEEEDSGSAAQLGCMNREIVR 1873  
QY 629 EGTISLPCQTFYQVCHETVGKFIQYGLTVAEHDDQEDISPSLASQQWDDKKLPEPLSWRS 688  
Db 1874 RGALILFCDYVCQNLHDS-----EHLTWLI 1898  
QY 689 DEEDEDSDGEGQDCYLYKVSQSKHQQTIFLQR---LLGPLLBAYSAALFVHNFSGP 745  
Db 1899 VNHIOQ-----LISLSHEPPVQDFISAIHRNSAASGLFIOAIRSC---ENLSTP 1945  
QY 746 VPEPEYLQKLKYLITRTRNVAVVAESATYCLVKNVAKMFKDIGVFETKOKRVSVLEL 805

Db 1946 TTLKKTQCLGHIHSQSGAVLTVYDRLLTSSRALARMV-DTLACRRVEMLLAANLQS 2004

Qy 806 SSTFLPQCNRQKLEYI 822

Db 2005 SMAQLPEEELNRIQEH 2021

RESULT 9

US-08-246-982A-16  
; Sequence 16, Application US/08246982A  
; Patent No. 5686288

GENERAL INFORMATION:

; APPLICANT: MacDonald, Marcy E.  
; APPLICANT: Ambrose, Christine M.  
; APPLICANT: Duyao, Mabel P.  
; APPLICANT: Gusella, James F.  
; TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/246,982A  
; FILING DATE: May 20, 1994  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

; NAME: Goldstein, Jorge, A.  
; REGISTRATION NUMBER: 29,021  
; REFERENCE/DOCKET NUMBER: 0609.3880002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3119 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-246-982A-16

Query Match

Best Local Similarity 2.7%; Score 115.5; DB 1; Length 3119;  
Matches 167; Conservative 145; Mismatches 343; Indels 303; Gaps 43;

Qy 12 DVSYLPHSEVSGVKHSTSEWGCGRPTVFRSATLKWESLMRKRPFVGRG----- 66

Db 1221 DVLKATHAYKVTLQDNSTKFG-GF---LRSALDVLQLELATLQDQKCVBEVL 1274

Qy 67 -----CYS-----CTPQSWDFKFNFSIPSL--GLRNVIYINETHRRHGLARLSY 111

Db 1275 GYLKSCFSREPMATVVCVQLLKTFLGTNLASQFDGLSS---NPSKQCR---AQLRG- 1326

Qy 112 VLFQIERDVHKGMF-----ATNVNVLNSSRVQEAIEVAELNPDG-----SA 156

Db 1327 -----SSSVRPLHYHCFMAYPTHQTQALASLRNMVQAE--QERDASGFVDVLQKVA 1379

Qy 157 QOQKAVNKVKKA-KRIQEMVATVSPAMIRLTGVLVLLKLFNSPFMNIQHKGOLEMVK 215

Db 1380 QLKNTLTSVTKNRADKNAIHNHRLFEPLVIK-----ALKQYTT-TTSVQLQKQVLDLLA 1433

Qy 216 AATENLPLFLPVHRSHIDVLLTF----- 241

Db 1434 QLVQLRVNVCLLDSQVDFIGFVLKQFIEVQGFRESEAIIPNFFFLVLLSYERYHSKQ 1493

Qy 242 -----ILFCHNTKAPYIASGNNL--NPIFSTLIHKLGFFIRRLDFTPOGRKDV 291

Db 1494 IIGIPKIIQLCDGI---MASGRKAVTHAIPALQPIVHDL---FVLRGTNKADAGELET 1546

Qy 292 YRALLGHIVELLRQOQFLFIEGTRSRSKTSCARAGLLSVVVDLTSTNVIPDILIP 351

Db 1547 QKEVVVSMRLRLIQYHQLVEMF-----ILVLQ 1573

Qy 352 VGISYDRIIEGHYNGEQLGKPKKNESLWSVARGVIRMLRKNYGCVRVDFAQPFSLKEYLE 411

Db 1574 -----QCH-----KENEKWK-----RLSRQ-----VADIILPLAKQOMH 1604

Qy 412 SQSQKPVSALESLEQALLPAILPRPSDADEGSDTTSINESRNATDSLR----- 461

Db 1605 IDSHEALGVNLTFEILAPSSL--RPVDM--LRSMFTPTSTMASVSTVQLWISGILAIL 1660

Qy 462 RRLTANLAELHILFTASKSCAI---MSTHIVACLLLYRHROG---IDL-----STLV 506

Db 1661 RVLISQSTEDIVL-----CRIQELSPHLLSCPVLNLRGGGNGVNLGECSEKQKSLP 1715

Qy 507 EDFF-----VMKEEVLARDFDLGFSGNSEPDVVMHAIQLLGNLCVVTITTSRNDP-F 556

Db 1716 EDTESRELLQVLGILLEDIVTKQKVDMSQOQHTFYCOELGTLLMC--LIHIFKSGMFR 1773

Qy 557 ITPSTT---VPSVPELNFYNGVLHVIME-----AIIACSLYAVLN----- 595

Db 1774 ITAAATRLFTSDGCEGSFYTLESINARVSRMVPHPALVLLWCQILLINHTDRHWAEV 1833

Qy 596 -----KRGLGPTSTPPNLISOELVRKAASL--CYL-LSNEGTSISLPQTFYQVCHETV 647

Db 1834 QQTPKRHSLSCTKSLNPKSGEEDSGSAQQLGCNREIVRRGALILFCDYVCQNLHDS- 1892

Qy 648 GKFIQYGLITVAEHDDQEDISPLAEQWDKCLPELSWSRDEDESDDFGEEQDCVLK 707

Db 1893 -----EHLTWLVINHIQD-----LIS 1908

Qy 708 VSQSKHQQTITLQF---LIGPLLEAYSSAAIEVHNFGFVPPPEYLQKLHLYITRE 764

Db 1909 LSHPEPPVQDFISAIHRNSAASGLFIQIQSRC---ENLSTPTTLKKTLCLEGIHLSQSG 1965

Qy 765 RNVAIYAESATYCLVKNVAMFKDIGVKETQKRVSVLELSSTFLPCNCRQKLEYI 822

Db 1966 AVLTLYVDRLGLTTPFRALARMV-DTLACRRVEMLLAANLQSSMAQLPEEELNRIQEH 2022

RESULT 10

US-08-453-265-16  
; Sequence 16, Application US/08453265  
; Patent No. 5693757

GENERAL INFORMATION:

; APPLICANT: MacDonald, Marcy E.  
; APPLICANT: Ambrose, Christine M.  
; APPLICANT: Duyao, Mabel P.  
; APPLICANT: Gusella, James F.  
; TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/453,265  
; FILING DATE: 30-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:



```
/ NAME: Ludwig, Steven R.
/ REGISTRATION NUMBER: 36,203
/ REFERENCE/DOCKET NUMBER: 0609.3880003
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 371-2600
/ TELEFAX: (202) 371-2540
/ INFORMATION FOR SEQ ID NO: 16:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3119 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-453-265-16

Query Match      2.7%; Score 115.5; DB 1; Length 3119;
Best Local Similarity 17.4%; Pred. No. 0.18; Mismatches 343; Indels 303; Gaps 43;
Matches 167; Conservative 145;

QY 12 DVSYLPHSSEYSVGRCKHTSEWGECCGFRPTVFRSATLKWESLMGRKRPFGVRC-----66
Db 1221 DVLKATHANYKVTLDLQNSTEKEG--GF-----LRSALDVLSQLLELATLQDICKVEEVL 1274

QY 67 -----CVS-----CTPSQWDKFNPSIPSL--GLRNVIYINETHRHGWLARLSY 111
Db 1275 GYLKSCFSREPMMATVCVQQLKTLFGTNLASQFDGLSS-----NPSKSQCR---AORLG- 1326

QY 112 VLFQERDVHKGME-----ATNVTENVLNSSRVQEAIAEVAELNPDG-----SA 156
Db 1327 -----SSSVRPGLYHYCFMAYPTHFTQALADASLRNNVQAE--QERDASGFVLOKVSA 1379

QY 157 QOQSKAVNKVKKA-KRILQEMVATSPAMIRLTGWLLKLFNFSFWNIQIHKGQLEWVK 215
Db 1380 QLKTNLTSTVTKRADKNAIHNLFEPLVIK-----ALKQYTT-TTSVQLQKQVLDLLA 1433

QY 216 AATETNLPPLFLPVHRSHIDYLLTF-----241
Db 1434 QLVQLRVNYCLLSDQVFIGFVLKQFEYIEVGOFRESEAIIPNIFPVLVLSYERYHSKQ 1493

QY 242 -----ILCHNIKAPIYASGNL---NIPFSTLIHKLGGFFIRRLDETDPGRKQVL 291
Db 1494 IIGPKIIQLCDGI-----MASGRKAVTHAIPALQPIVHDL---FVLGRTKNADAGELET 1546

QY 292 YRALLGHGIVELLRQOQFLBIFLEGTRSRGKTSARAGLLSVVVDLTSTNVPIDILIIP 351
Db 1547 QKEVVSMRLRLQYHQVLEMF-----ILVLQ 1573

QY 352 VGISYDRIIEGHYNGEOLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPPSLKEYLE 411
Db 1574 -----QCH-----KENEDKW-----RLSRQ-----VADILPLMAKQOMH 1604

QY 412 SQSQKPSALLSLEQALLPAILSRPSDADEGRDTSINESRNATDESRL-----461
Db 1605 IDSHEALGVLTLEFILAPESSL--RPVDMIL--LRSMFITESTWASVSTVQLMISGILAIL 1660

QY 462 RRIIANLAHEILFASKSCAI-----NSTHIVACLLYRHROG---IDL-----STLV 506
Db 1661 RVLISQSTEDIVL-----CHIQELSPHLLSCPVINLRGGGNTVLEGCSGKQKSLP 1715

QY 507 EDFP-----VMKEVILARDFDLFGSGNSEDVVMHAIQLLGNVCVTIHTSRNDEF-F 556
Db 1716 EDTFSRPLLQVGLLEDIVTKQKVDMSQQHTFYCOELGTLLMC--LIHIFKSGMFR 1773

QY 557 ITPSTT---VPSVFELNFYNGVLHVIME-----AIIACSLYAVLN-----595
Db 1774 ITAAATRLFTSDCEGSFYTLESNARVRSMVPTHPALVLLWCQILLNHTDHRWAEV 1833

QY 596 -----KRGGLGPTSTPNLISQOLVKRAASL--CYL-LSNEGFTISLPCQTFYQVCHETV 647
Db 1834 QOTPKRHSLSCTSLNPKQSGEEDSGSAQLGNCREIVYRGALILFCDYVCQNLDHS- 1892

QY 648 GKFIQVIGILTVAEHDDQEDISPSLAEQQWDKKLPEPLSWRSDEDESDFGEEQRDCYLK 707
Db 1893 -----EHLTWLIVNHQID-----LIS 1908
```

---

```
QY 708 VSQSKHQQFIFLOR---LLGLLEAYSSAAIFVHNFSGPVPEPEYLOKLHKYLITRTE 764
Db 1909 LSHEPPVQDFISAIHRNSAASGLFQAIQSRC---ENLSTPTTLTKTLCGLSIHLSQSC 1965

QY 765 RNVAVYASATYCLVKNAVNFKDIGVFKBTQKRVSVLELSSTFLPQCNROKLLBYI 822
Db 1966 AVTLTYVDRLLGTTPFRALARMV-DTLACRRVEMLLAANLQSSMAQLPEBELNRIOEHL 2022

RESULT 11
US-09-315-793-12
; Sequence 12, Application US/09315793
; Patent No. 6221597
; GENERAL INFORMATION:
; APPLICANT: Roberts, Christopher J.
; TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL
; TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATION
; FILE OF INVENTION: DRUGS
; FILE REFERENCE: 9301-048
; CURRENT APPLICATION NUMBER: US/09/315,793
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1698
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-315-793-12

Query Match      2.6%; Score 109.5; DB 3; Length 1698;
Best Local Similarity 18.2%; Pred. No. 0.26;
Matches 163; Conservative 137; Mismatches 289; Indels 309; Gaps 44;

QY 75 WDKFNPSPISGLRNVIYINETHRHGWLARLSYLVFIQERDVHKGMEFATNTVENVL 134
Db 706 WKSLKDPRHQILGVKNLKTNR-----FIPSKFTESALVVTFVEE---745

QY 135 NSRVQEAIAEVAELNPDGSAQOQSKAVNKVKKAKRILQEMVATSPAMIRLTGWVL- 193
Db 746 -----EDISERLSVL--DLLWTQLDSDSNLRRPLELILGELFDDQNPFFYLTVSKWILS 797

QY 194 -----LKLNFNSFWNI---QIHKGQLEWVKAATETNLPPLFLPVHRSHIDYLLTFI 242
Db 798 ILNSGSASRLFYLTNTILNKNRLEKRLD-----ERDLDMLTYEFQ 840

QY 243 LFCHNIKAPIYASGNLNIPIFSTLIHKLGGFFIRRLDETDPGRKQVLYRALLHGHIVE 302
Db 841 MLAYVLK-----TNGRTRKVFSTELTSIKSTIKWNEDVS-----TYKSL---LVT 885

QY 303 LLRQOQFLBIFLEGTRSRGKTSARAGLLSVVVDLTSTNVPIDILIIPVGISYDRIIEG 362
Db 886 LMR---FLNI-KSNTHAKSIRSALI---LLDILLDGTQCN-FKDIVIFLLQWSSKYIABE 937

QY 363 HYNGEOLGPKKNESLWSVARGVIRMLRKN-----YGCVRVDF-----400
Db 938 GIEPELIA-----VSLDIVSKVLRSLSHNGIKLIDFDDNAAHLKY---IDFLVTSVN 988

QY 401 -AQPSLKEYLSQSQKPSALLSLEQALLPAILSRPSDADEGRDTSINESRNATDES 459
Db 989 MKSPLIIVAYVKLLSE---SIVYFENSIFRMILP-----1019

QY 460 LRRRIIANLAHEI--LFTASK-----SCAIMSTHIVACL-----LLYRHROG-IDLS 503
Db 1020 -----LSASLVQCVRQLFLLKEBEGGYQPIALLLGLLEELLEISHGYLVEEREGVFSG 1075

QY 504 TLVEOPFVMKEVILARDFDLFGSGNSE-----DVVMHAI-QLLGNVCVTIHTSR-551
Db 1076 NLKGFPI---QSVSNVFS-SDSSEESKIQERDVILQSFQVISCCLDIWYWAHNISC 1131

QY 552 ---NDEFFITPSTVPSVFELNFYNGVLH-VFIMEAI-----IACSLYAVL-594
Db 1132 KSNDDSSL--DATNHSYKFKFRSKKLETLFLLBPELLENLISRSNTTTLVHVHLD 1189
```

```

Qy 595 -NKGGLGPTSTPPNLI-----SQBLVRKAASLCYLLSNEG 630
      ||      ||      ||      ||      ||      ||      ||
Db 1190 GNK-----PAITIPHLLYGVIIYRNRTASVFKFSNRDGRSSTTKLTKGSPMLKRLSGES 1244
      ||      ||      ||      ||      ||      ||      ||
Qy 631 TISL-----PCQTFYOVCHETVKCFQYIGILTVAEHDDQEDISPSLAEOQWDKK 679
      ||      ||      ||      ||      ||      ||      ||
Db 1245 IIAELFNLYVDSVENSAMEEFY-----GDFLLFFREVATNLYNLDVLSIL-----K 1291
      ||      ||      ||      ||      ||      ||      ||
Qy 680 LPEPLSWRSDDEDESDFGEEOR-----DCVLK----- 707
      ||      ||      ||      ||      ||      ||      ||
Db 1292 LVALLSGVSK-----TFGEQKRVREISDVFFKYLPNAPINFNTNLYRGHPDSFKOLEFV 1347
      ||      ||      ||      ||      ||      ||      ||
Qy 708 -----VSQSKHQOQFTFLQRLGPLLEAYSSAAI--FVHNFSGPVEPEYLQK 754
      ||      ||      ||      ||      ||      ||      ||
Db 1348 VWRVQYIVNDQIGGPKFNTLTIIVNOCLTPPIKPKSEKTIPOYVULEAAVSHLGSKV 1407
      ||      ||      ||      ||      ||      ||      ||
Qy 755 LHKYLIT---RTERNAVVAESATYCLVKVNAFMKMDIGVFKEKTKRVS--VLELSS 807
      ||      ||      ||      ||      ||      ||      ||
Db 1408 SWRLLJIAELFQNDKKLSVIGSDQTW-----EKIIYEWISIYPENKSKILNDLLLEIGS 1459
      ||      ||      ||      ||      ||      ||      ||

RESULT 12
US-09-328-352-7763
; Sequence 7763, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 763
; LENGTH: 2188
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-7763

```

Query Match	2.5%;	Score 105.5;	DB 4;	Length 2188;	
Best Local Similarity	18.8%;	Pred. No. 1.1;			
Matches 113;	Conservative 98;	Mismatches 220;	Indels 169;	Gaps 27;	
QY	272	GGFIRRLDET	PDGRKDV--LYRALLHGHI	VELLRQOQFLEIFUEGRSRSGKTSC----	326
DB	1115	GKFDTELEDK	FABQLRGVNQYS	NOVDPELLDYSQDMDLQRYAEG--RTYKGKLSLDWIE	1173
QY	327	ARAGLLSVVDTL	STNVI-----PDLLIIPGI	-SYDRIIEGHYNQEOLGKPKKNE	376
DB	1174	AKYGRDSIDI	QISIKGAYGFAQR	GENPDVVAEMFGYESGDALRDLNLSN	P--SPKQKI 1230
QY	377	SLWSVARGVIR	MLRKNYGCVRDV	FAQPFSLKEYLSEOSQKPVSA--LLSLRQALLPATLPS	435
DB	1231	DELTDARMAVQY	-----SEFFDOOSII	EAVEAALHNDVRARMLSAEWAALNGLT--	1279
QY	436	RPSDAADEGR	TSINE--SRNATDSLR	RRRLIANLAELHILFASKSCAIMSTHIVACILLY	494
DB	1280	-----GRKSALNE	AAKTVAQDIVOR	KIKDIRPHV-----	1309
QY	495	RHROGIDLS	TIVEDFVYMKREV	-----LARDFDLPGSGNSEDVVMHAIQALL	-----G 541
DB	1310	RAQDDAELGR	MANEAFKGETVEA	ARHNRNQLVGFYATKYSYDAKQOTQKHLDLVKKVFG	1369
QY	542	NCVTITHTSR	NDEF--FITPSTTV	PSVFPFLFYNSGVNLHFVIMEAIIACSILYAVLNKRGGLG	600
DB	1370	NNEKL-----SK	RDPDFVTAARGIL	GKYGDLGRESTNYEHOELIRKYDPTTYAETQNI	G-- 1424
QY	601	GPTSTPPN	-----LISQELVR	KAAASLYLLSNEGTISLPCQTFYOVCHETVCKFTQYG	654
DB	1425	---ALPENQNV	REULTQOFQNAV	MAAVETLMHRSKEN-----KIWHITTEAF	--- 1467
QY	655	ILTVAEHDDO	EDI	SPSLAEQOWDKKLPELISWSRDEDEDSDFGEORDCVLKVSQS--KE	713

```

Db      1468  -----EREQREELIQOTGGKKSVKEIQOTLLGRDKTA-----ELUKAFMELGASAKR 1515L
Qy      714  HQQFITFLQ-----RLLGPLLEAYSSAAIFVHNFGVPPEYLOKLHKYLIIRTE 764
Db      1516  VDQVITWLDGGANGKFRTYLINPMQDA-----LAKYRIEKAK 1552L
Qy      765  RNVAVYAESATYCLVKNVAVNMFNDIGVFKETKOKRVSVLELSS--TFLPQCNROKLLBYIL 823
Db      1553  -----MLKDDVDVIFEGFGKLDNSK---IAAPELNNTFTV---GKQSLHAIL 1593L

RESULT 13.
US-09-770-170-6
; Sequence 6, Application US/09770170
; Patent No. 6319679
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven
; APPLICANT: Gardner, Kevin
; APPLICANT: Harper, Shannon
; APPLICANT: Rutter, Jared
; APPLICANT: Michnoff, Carolyn
; APPLICANT: Amezcua, Carlos
; TITLE OF INVENTION: PAS Kinase
; FILE REFERENCE: UTSD:0761
; CURRENT APPLICATION NUMBER: US/09/770,170
; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 1356
; TYPE: PRt
; ORGANISM: yeast
US-09-770-170-6

```

```

Query Match      2.4%; Score 103; DB 4; Length 1356;
Best Local Similarity 19.0%; Pred. No. 0.84;
Matches 89; Conservative 82; Mismatches 173; Indels 124; Gaps 21;

QY      297 HGHIVELLRQQQFLEIFLECTRSRGKTSARAGILLSVVVDTLSTNV--IPDILLIPVG 353
Db      324 HGPSMDILNEQ-----RANLKS-----LNLNETLEKNTSERASDLHMISL- 365
QY      354 ISYDRIIEHYNGEQLGPKPKNESLVSARGVIRMLRKKNYGCVRVDPQAQPSLKEYLESQ 413
Db      366 -----FWNLKMLGDPKKNNSERDKRTEKKKILLD-----SLAEFF--EHNFI 409
QY      414 SQKPVGALLSLEQALLPAILSPSPDAADRGRTSINESRNATDESRLRLIA--NLAE 470
Db      410 GDNPIADTDLKEEI-----DEFTGSDTATIDIRPOODYGRILFTFTKNSAP 460
QY      471 HILFTASKS--CAIMSTHIVACLLLYRHRGIDILSTLVEDF-----FVMKEVLARDFD 522
Db      461 QAIPTCSQDPWQFRAANDLACLVLFGISQNAIRALTMLDLHTDSRNFVLKLLSTEGQE 520
QY      523 LGFSGNSDVMMAIQLLGNCVTITHTSRNDFFITPSTTVPS-VFELNYS---NGVLH 578
Db      521 MVFTG-----EIIIG-----IVQPETLSSSKVWVASFWAKRKNGLL- 555
QY      579 VFTMEAILACSLYAVLNKRGLGFTSTPPNLIISQQLVKAASLCVLLSNEGTLISLPCQT 638
Db      556 VCVFEKVPQDYDVLNLNDDFGA---ENIVDKCELLSDGPT----LSSSLSLSP--- 603
QY      639 FYQVCHETWKGFTQYGLITVAEHDDQEDISPSLAEQQWKKLPEPLSWRSDEDESDDFG 698
Db      604 --KXASPTSGKLEY-----SLERKILEKSYKPTSTENRNGDENOLDG 645
QY      699 EQQRDCYLKVS--QSKEHOOF-----ITFLORLLGPLLEAYSAALF 738
Db      646 DSHSEPSLSSSPVRSKSKVKFANDIKDVKISOSLAKLMDMDVRNGVVF 693

```

US-09-252-991A-24404  
; Sequence 24404, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24404  
; LENGTH: 531  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24404

Query Match 2.4%; Score 102; DB 4; Length 531;  
Best Local Similarity 21.9%; Pred. No. 0.2;  
Matches 61; Conservative 41; Mismatches 106; Indels 70; Gaps 11;  
QY 179 ATVSPAMIRLTGWLKLFNSFFWNQIHKQGLEMKAAATETNLPLFLPVHRSHIDYLL 238  
DB 309 ADVTLRQRLTRWLARLCAALFEVRV-----SGEAPROPMLVANHVSWTDIPL 359  
QY 239 LTFILFCHNIKAP--YIAGNNINIPFSTLIHLKLGFFIRRLDETTPDGRKDVLYRALL 296  
DB 360 LGAL-----APLTFLSKAEVRAWPLAGLAEKAGTLYIRR--GSGDSR-----LI 402  
QY 297 HGHIVELLRQOQLEIFLEGTGSRSGKTSRAGLLSVVDLTSTNVIPDILLIPVGISY 356  
DB 403 NORLAEQLHGRNLLIFPEGTITNGESLRTFHGLMASALEA-----GVAVQPVVAISY 455  
QY 357 DR-----IIEGHYNGE-----QLCKPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSL- 406  
DB 456 RRGVDPDTQAPFGIDDDLLSHLG-----RLLRGERGSHVHQLLEPIPSQ 499  
QY 407 ---KEYLESQKPVSAALLSLEQALLPAILPSRPSDAA 441  
DB 500 GLDRAELARQAQAV-----RLALFGTAAPTQTRRAA 531

## RESULT 15

US-09-417-822-2  
; Sequence 2, Application US/09417822  
; Patent No. 6344549  
; GENERAL INFORMATION:  
; APPLICANT: Keegan, Kathy  
; TITLE OF INVENTION: ATR-2  
; FILE REFERENCE: 27866/35633  
; CURRENT APPLICATION NUMBER: US/09/417,822  
; CURRENT FILING DATE: 1999-10-14  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 2930  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-417-822-2

Query Match 2.3%; Score 100.5; DB 4; Length 2930;  
Best Local Similarity 19.5%; Pred. No. 6;  
Matches 137; Conservative 102; Mismatches 254; Indels 209; Gaps 34;  
QY 218 TETNPLLELPVHRSHIDYLLTFLFCHNIKAPYIAGNNINIPFSTLIHLKLGFFIR 277  
DB 14 SETYALFLSLP-----SFHKFKGLLANTVEDVNICIQACSSL-HALS----- 56  
QY 278 RLDETDPGRKDVLYRALLHGHIVELLRQOQLEIFLEGTGSRSGKTSRAGLLSVV 335

DB 57 ---SSLPD---DILQRC-----VDVCR-----VQLVHSGTRIRQAFGL- LKSIIP 94  
QY 336 VD-TLSTNVIPDILLIPVG-----ISVDRIIEGHYNGEQLGKPK 373  
DB 95 LDVVLNNNHTEIQEISLALRSHMSKAPNTTHPQDFSDVISF--LYG--NSHRTGKON 150  
QY 374 KNESLWSVARGVIRMLRKNYGCVRV-----YSCQRLDKRQSTIPRNLKTDVLMQWAIWEAAQFTVLSK 409  
DB 151 WLERLF-----YSCQRLDKRQSTIPRNLKTDVLMQWAIWEAAQFTVLSK 197  
QY 410 LESQSQKPVSAALLSLEQALLPAILPSRPSDADEGRDTSINERSNATDSLRRLRLIANLA 469  
DB 198 LRTPLGRAQDTTQTIE-GIIRSLAAHTLNPQDVQSWTTADNDEGHGNNQLRLVLLQYL 256  
QY 470 EH---ILFTASKSC--AINSTHIVACLLLYRHRQGI-DLSTLVEDDFVMKEEVLARDFDL 523  
DB 257 ENLEKLMYNAYEGCANALTSPPKVIPTFYTNROTQDWLTIR-LSIMRVGLLA----- 310  
QY 524 GFSGNSDVVMEAIQLLGNCVITITHSRNDFFITPSTTVPSVFELN-----FYSN 574  
DB 311 ---GQPAVTVRHGFDLLTEMKT-TLSLQNELEVTIMVVEALCELHCPEAQIAGVMS 366  
QY 575 GVL-----HVFIMEAIIAC-SLY--AVLNKRGLGGPTS 604  
DB 367 SIYGNLLWINSVAQQAEGREFEKASVEYQEHLCAMTGVDCISSFDKSVLTLANAGRNA 426  
QY 605 TPNLISQQLYRKAASLCYLLSNEGTISLPQCTFYOVCHETVKGFIQYIGILTVAEHDDQ 664  
DB 427 SPKHSNGE-----SRKTVLSKPTDSSPEVINYLGNKACEYI----- 464  
QY 665 EDISPLAQOWDKKLP-PLSWRSEDEDESDFGEEQRDCYLKVSQSKEHQOQFITFLOR 723  
DB 465 -SIADWAAVOEQWNAIHDLLKSTSTSLNLKADFN-----YKLSLSFESGKFEVCTEQ 517  
QY 724 L-LGP-----LLEAVSSAAIFVHNFSGPV--PEPEYLOKLHKYLIIRTRERNVA-----VY 770  
DB 518 LELLPGENINLAGGSKKIDMKLLPNMSPDPRELOKSIEVQLLRSSVCLATALNPTE 577  
QY 771 ABSATYCLVKNVAKFKDIGVFKEYKQKRVSVLELSSTLPQ 812  
DB 578 QDQKQOSITENVVK-----YLKQTSRIAIGPLRLSTLTVSQ 613

Search completed: January 10, 2004, 20:22:33

Job time : 32 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2004, 20:20:58 ; Search time 58 Seconds  
(without alignments)  
2876.725 Million cell updates/sec

Title: US-09-935-290-2

Perfect score: 4280

Sequence: 1 MDESALTGTIDVSYLPHSS.....FLPQCNRQKLEYILSFVVL 828

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4280	100.0	828	11	US-09-935-290-2
2	4279	100.0	828	12	US-10-288-252-6
3	936.5	23.3	801	12	US-10-288-252-7
4	928.5	21.7	775	12	US-10-353-690-32
5	678	15.8	718	12	US-10-369-493-6322
6	505	11.8	870	12	US-10-369-493-17551
7	500	11.7	754	12	US-10-369-493-9385
8	490	11.4	833	12	US-10-369-493-13643
9	486	11.4	754	12	US-10-369-493-9240
10	452	10.6	772	12	US-10-369-493-16014
11	452	10.6	792	12	US-10-369-493-15639
12	452	10.6	795	12	US-10-369-493-16367
13	407.5	9.5	698	12	US-10-369-493-13620
14	245	5.7	789	10	US-09-712-363-244
15	128	3.0	1095	12	US-10-094-749-1905

16	119.5	2.8	1174	15	US-10-205-841-40	Sequence 40, Appl
17	117.5	2.7	195	9	US-09-817-910-11	Sequence 11, Appl
18	117.5	2.7	195	9	US-09-798-029-27	Sequence 27, Appl
19	117.5	2.7	195	9	US-09-887-389-4	Sequence 4, Appl
20	117.5	2.7	195	10	US-09-924-358-46	Sequence 46, Appl
21	117.5	2.7	195	10	US-09-882-872-4	Sequence 4, Appl
22	114	2.7	714	12	US-10-369-493-10138	Sequence 10138, A
23	111	2.6	987	12	US-10-369-493-21998	Sequence 21998, A
24	108	2.5	788	12	US-10-369-493-11178	Sequence 11178, A
25	108	2.5	1420	12	US-10-310-154-387	Sequence 387, App
26	108	2.5	3529	12	US-10-165-216-8	Sequence 8, Appl
27	108	2.5	3657	12	US-10-165-216-10	Sequence 10, Appl
28	106	2.5	532	12	US-10-104-047-2555	Sequence 2555, Ap
29	103	2.4	432	15	US-10-220-380-4	Sequence 4, Appl
30	103	2.4	1356	15	US-10-059-963-6	Sequence 6, Appl
31	103	2.4	3031	12	US-10-165-216-6	Sequence 6, Appl
32	103	2.4	3521	12	US-10-165-216-2	Sequence 2, Appl
33	102.5	2.4	1847	12	US-10-032-585-7538	Sequence 7538, Ap
34	102	2.4	454	11	US-09-963-339-2	Sequence 51, Appl
35	102	2.4	454	12	US-10-145-586-51	Sequence 22419, A
36	102	2.4	700	12	US-10-369-493-22419	Sequence 21903, A
37	102	2.4	805	12	US-10-369-493-4333	Sequence 433, App
38	102	2.4	4746	12	US-10-369-493-22294	Sequence 22294, A
39	101	2.4	1661	12	US-10-369-493-19598	Sequence 19598, A
40	100.5	2.3	615	12	US-10-369-493-22294	Sequence 2, Appl
41	100.5	2.3	2930	11	US-09-957-837A-2	Sequence 3180, Ap
42	100	2.3	463	12	US-10-094-749-3180	Sequence 5843, Ap
43	99	2.3	550	10	US-09-738-626-5843	Sequence 20, Appl
44	99	2.3	4351	12	US-09-970-944-20	Sequence 8, Appl
45	98.5	2.3	872	10	US-09-843-676-8	

ALIGNMENTS

RESULT 1

US-09-935-290-2  
; Sequence 2, Application US/09935290  
; Publication No. US20030044948A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller-Libermann, Rosana  
; TITLE OF INVENTION: 56919, A NOVEL HUMAN ACYLTRANSFERASE AND USES THEREOF  
; FILE REFERENCE: MNI-186  
; CURRENT APPLICATION NUMBER: US/09/935,290  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/226,509  
; PRIOR FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 828  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-935-290-2

Query Match		100.0%;	Score 4280;	DB 11;	Length 828;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 828;		Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	MDESALTGTIDVSYLPHSSEYSGRCCKHTSEEWGCGFRPTVFRSATLKWESLMSGRK	60		
Db	1	MDESALTGTIDVSYLPHSSEYSGRCCKHTSEEWGCGFRPTVFRSATLKWESLMSGRK	60		
Qy	61	PFVGRCCYCTPQSDWDFNPSLGLRNVIIYNETHRHGWLARLSVYLFIQERDV	120		
Db	61	PFVGRCCYCTPQSDWDFNPSLGLRNVIIYNETHRHGWLARLSVYLFIQERDV	120		
Qy	121	HKGWFATNVNTVNLSSRVQEAIAEVAELNPDGSAQQSKAVNKKKAKRIIQEWVAT	180		
Db	121	HKGWFATNVNTVNLSSRVQEAIAEVAELNPDGSAQQSKAVNKKKAKRIIQEWVAT	180		
Qy	181	VSPAMIRLTGVLVLLKLFNSFFWNIQIHKGQLEWKAATETNLPLLFLPVHRSHDYLILT	240		
Db	181	VSPAMIRLTGVLVLLKLFNSFFWNIQIHKGQLEWKAATETNLPLLFLPVHRSHDYLILT	240		

Db 181 VSPAMIRLTGWVLLKLFNSFFWNIQIHKGQLEWVKAATNTLPLLPVHRSHIDYLLLT 240  
Qy 241 FILFCHNIKAPYIASGNLNIPIFSTLIHKLGGFFIRRRDETDPDGRKDVLYRALLHGI 300  
Db 241 FILFCHNIKAPYIASGNLNIPIFSTLIHKLGGFFIRRRDETDPDGRKDVLYRALLHGI 300  
Qy 301 VELLROQOFLEIFLEGTRSRSGKTSARAGLLSVVDTLSTNVIPDILIPVGISYDRII 360  
Db 301 VELLROQOFLEIFLEGTRSRSGKTSARAGLLSVVDTLSTNVIPDILIPVGISYDRII 360  
Qy 361 EGHNGEOLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLESQKPVSA 420  
Db 361 EGHNGEOLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLESQKPVSA 420  
Qy 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
Db 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
Qy 481 AIMSTHIVACLLYRHRQGDISTLVEDFFVMKEEVLARDFDLFGSGNSEDVVMHAIQLL 540  
Db 481 AIMSTHIVACLLYRHRQGDISTLVEDFFVMKEEVLARDFDLFGSGNSEDVVMHAIQLL 540  
Qy 541 GNCVTIHTSRNDEFFITPTSTVPSPFELNFYSGVLHVIMEAIIACSLYAVLNKRGIG 600  
Db 541 GNCVTIHTSRNDEFFITPTSTVPSPFELNFYSGVLHVIMEAIIACSLYAVLNKRGIG 600  
Qy 601 GPTSTPNLISQQLVRKAASCLYLSNEGTSILPCQTFYQVCHETVKGFIQXGILTVAE 660  
Db 601 GPTSTPNLISQQLVRKAASCLYLSNEGTSILPCQTFYQVCHETVKGFIQXGILTVAE 660  
Qy 661 HDQEDISPSPSLAEQWQDKKLPEPLSWRSDEDESDFGGEQRDCYLKVSQKEHQOFTIF 720  
Db 661 HDQEDISPSPSLAEQWQDKKLPEPLSWRSDEDESDFGGEQRDCYLKVSQKEHQOFTIF 720  
Qy 721 LQRLGLPLLEYASAAIFVNFSGPVPPEPYLOKLHKYLIITRERNVAVVAESATYCLVK 780  
Db 721 LQRLGLPLLEYASAAIFVNFSGPVPPEPYLOKLHKYLIITRERNVAVVAESATYCLVK 780  
Qy 781 NAVMFKDVGFKETKQKRVSVLELSSTLPQCNROKLLYIISFVVL 828  
Db 781 NAVMFKDVGFKETKQKRVSVLELSSTLPQCNROKLLYIISFVVL 828

## RESULT 2

US-10-288-252-6  
; Sequence 6, Application US/10288252  
; Publication No. US20030143686A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: LAL, Preeti G.  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: YUE, Henry  
; APPLICANT: BURFORD, Neil  
; APPLICANT: GANDHI, Aameena R.  
; APPLICANT: WARREN, Bridget A.  
; APPLICANT: YAO, Monique G.  
; APPLICANT: TRIBOULEY, Catherine M.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: LEE, Ernestine A.  
; APPLICANT: HAPALIA, April J.A.  
; APPLICANT: LU, Yan  
; APPLICANT: GRIFFIN, Jennifer A.  
; APPLICANT: SANJANWALA, Madhu S.  
; APPLICANT: DING, Li  
; TITLE OF INVENTION: TRANSFERASES  
; FILE REFERENCE: PI-0241 USA  
; CURRENT APPLICATION NUMBER: US/10/288,252  
; CURRENT FILING DATE: 2002-11-04  
; PRIOR APPLICATION NUMBER: PCT US01/30424  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 60/252,819  
; PRIOR FILING DATE: 2000-11-21  
; PRIOR APPLICATION NUMBER: US 60/249,639

; PRIOR FILING DATE: 2000-11-16  
; PRIOR APPLICATION NUMBER: US 60/247,931  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: US 60/246,001  
; PRIOR FILING DATE: 2000-11-03  
; PRIOR APPLICATION NUMBER: US 60/244,025  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: US 60/238,481  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/236,523  
; PRIOR FILING DATE: 2000-09-29  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PERL Program  
; SEQ ID NO 6  
; LENGTH: 828  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030143686A1 1517230CD1  
US-10-288-252-6

Query Match 100.0%; Score 4279; DB 12; Length 828;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 827; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDSEALTGTIDVSYLPHSSEYSGRCCKHTSSEWGECEGPRPTVFRSATLKWESLSMRKR 60  
Db 1 MDSEALTGTIDVSYLPHSSEYSGRCCKHTSSEWGECEGPRPTVFRSATLKWESLSMRKR 60  
Qy 61 PFVGRCCYCTQSQWMDKFFNPSPISGLRNVIYINETHTRHGWLARLSYVLFQERDV 120  
Db 61 PFVGRCCYCTQSQWMDKFFNPSPISGLRNVIYINETHTRHGWLARLSYVLFQERDV 120  
Qy 121 HKGMFATNTENVLNSRRVQEAIAEVAELNPDGSAQQSKAVNKVKKAKILQEMVAT 180  
Db 121 HKGMFATNTENVLNSRRVQEAIAEVAELNPDGSAQQSKAVNKVKKAKILQEMVAT 180  
Qy 181 VSPAMIRLTGWVLLKLFNSFFWNIQIHKGQLEWVKAATNTLPLLPVHRSHIDYLLLT 240  
Db 181 VSPAMIRLTGWVLLKLFNSFFWNIQIHKGQLEWVKAATNTLPLLPVHRSHIDYLLLT 240  
Qy 241 FILFCHNIKAPYIASGNLNIPIFSTLIHKLGGFFIRRRDETDPDGRKDVLYRALLHGI 300  
Db 241 FILFCHNIKAPYIASGNLNIPIFSTLIHKLGGFFIRRRDETDPDGRKDVLYRALLHGI 300  
Qy 301 VELLROQOFLEIFLEGTRSRSGKTSARAGLLSVVDTLSTNVIPDILIPVGISYDRII 360  
Db 301 VELLROQOFLEIFLEGTRSRSGKTSARAGLLSVVDTLSTNVIPDILIPVGISYDRII 360  
Qy 361 EGHNGEOLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLESQKPVSA 420  
Db 361 EGHNGEOLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQFSLKEYLESQKPVSA 420  
Qy 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
Db 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
Qy 481 AIMSTHIVACLLYRHRQGDISTLVEDFFVMKEEVLARDFDLFGSGNSEDVVMHAIQLL 540  
Db 481 AIMSTHIVACLLYRHRQGDISTLVEDFFVMKEEVLARDFDLFGSGNSEDVVMHAIQLL 540  
Qy 541 GNCVTIHTSRNDEFFITPTSTVPSPFELNFYSGVLHVIMEAIIACSLYAVLNKRGIG 600  
Db 541 GNCVTIHTSRNDEFFITPTSTVPSPFELNFYSGVLHVIMEAIIACSLYAVLNKRGIG 600  
Qy 601 GPTSTPNLISQQLVRKAASCLYLSNEGTSILPCQTFYQVCHETVKGFIQXGILTVAE 660  
Db 601 GPTSTPNLISQQLVRKAASCLYLSNEGTSILPCQTFYQVCHETVKGFIQXGILTVAE 660  
Qy 661 HDQEDISPSPSLAEQWQDKKLPEPLSWRSDEDESDFGGEQRDCYLKVSQKEHQOFTIF 720  
Db 661 HDQEDISPSPSLAEQWQDKKLPEPLSWRSDEDESDFGGEQRDCYLKVSQKEHQOFTIF 720

1



```
Db 407 AQMCSIQPVAVVSCLLAKMGKVSRSFTFEDCEWLCEKIIAEGDVGVGQSKTKGSA 466
Qy 532 VMHAIQLGNCVTITHTSRNDEFTPTSTVPVSFELNFYSNGVLHVFIMEAIIACSLY 591
Db 467 LVKYAFKLESCEVT-----DE-YVSPKSHSSFTITLAYKNKSNVICFISKSVALTI- 519
Qy 582 AVLNRKGLGGTSTPPNLIQSOBLVRKAASLCYLLSNGTISLPQCTFYQVCHETVGVKFI 651
Db 520 -----VSRPSGTK---LSIDQIVEDALSCLDWLQFBEFMCPCDSRLRELHNVLG--- 566
Qy 652 QXGILTVAEHDDQEDISPSLAEOQWDKKLPEPLS--WESDEDED--SDFGEEQRCYLKV 708
Db 567 -----QKENS-----DPIHGFLKSEIEDDGLDAGGALNSGTURV 601
Qy 709 SOSKEHQQFITFLQRLGLLGLLEAYSAAFVHNFSGPVPEPEYLOKLHKYLIITRT----- 763
Db 602 RDAKS-RETLOFFANLVRPFVQSLVLISSFVVEKCP-EPSTDNIIIRQLCQOSLAGDI 659
Qy 764 ERNVAIVAESATYCLVKNAVAMFKDIGVFKETKQKRV-----VLELSSTFLPQCNRQK 817
Db 660 DLPFAPLESINSDSFKNALRVLDKGLLQSTPNSTARGNSRLAELIS-----NLER 713
Qy 818 LLE 820
Db 714 VLE 716

RESULT 6
US-10-369-493-17551
; Sequence 17551, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17551
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-369-493-17551

Query Match 11.8%; Score 505; DB 12; Length 870;
Best Local Similarity 22.4%; Pred. No. 5.1e-41;
Matches 197; Conservative 148; Mismatches 314; Indels 222; Gaps 26;

Qy 13 VSYLPHSSEYSVGR-CKHTSEEWGEC-----GPRPTVFRSATLKWKESLMSRKPFGV 64
Db 64 VQYEHCVQYNDGRPIIVLEDEYGLCNTLIIDKACRKTLPSPILPNCNPLQKRAVLA 123
Qy 65 RCCVST-----PQSWDKFNPSPSL-----GLRNV-----IYINETHRRHGLA 106
Db 124 LSRSSNSLIPNQRGGKTHSDSLANLLQAHRIKRDLDVHLVPVSIFIGRTPDQSGWFA 183
Qy 107 -----RRLSYVLFIQ 116
Db 184 VLFSENWALVGRFRLLAVLNGRNTIVCFAPPI-SVRQTLNEGLPPTLRLKQRLRAH 243
Qy 117 ERDVHKGMAFNVT-----ENVLNRSRVQEAIAEVAELPDGSAQOOSKAVNVKKA 170
Db 244 FRRIRETVIGDPLSTRLLLVNDVATEAVRAIA-----SOAKRKGDTLSETWRKA 294
Qy 171 KRILQEMVATVSPAMIRLTGWLLKLFNSFFWNIQIHKQGLQEMVKAATETNLPLFLPVH 230
```

```
Db 295 QAYAWETAAADYSSPVIRSADFLFSGHVNRIYAGVLIH--HVDSEFK-ETAPGHEVVYVPSH 351
Qy 231 RSHIDYLLLTILFCHNIKAPYIAGNGLNIPITSLIHLKLGFFIRRLDETDPGRKDV 290
Db 352 RSHIDYLLLSYCLTQCGVILPHIVAGINLNFIVGTLLRKCGAFFIRSI-----KGNM 405
Qy 291 LYRALLHGHIVELLARQQQFLIEFLEGTSRSGTKSCARAGLLSVVVDTLSTNVIPDILII 350
Db 406 LYSIVLSEYVAQLVAGVSLSEYFIEGGRSRTGRLLPQKGGIMMTLQAFLAQRPRLVLFQ 465
Qy 351 PVGISYDRITIEG-HYNGEQLGKPKNESLWSVARGVIRMLRKNYGCVRVDPAQFSLKEY 409
Db 466 PIYIGYKLTGTSYLDLSGEPKKESIWRLFNWIPKVLKQKYGVVNVFGEPIALNDV 525
Qy 410 LESQSPVPSALLSLEQALLPAILPSRPSDAADGRDTSINESRNATDESRLRLIANLA 469
Db 526 L-----AELAPWEQOALNENEKPAWLS-----NTVNHLA 555
Qy 470 EHILFTASKCAIMSTHIVACLLYHRQGDIDSLTVEDEFVWKEEVLARDDFDLGFSG-- 527
Db 556 RQIOTRINSAADVNPINLLALALLSTPKHAMEADLIAQITLCKKILL-----ELPYSNRV 611
Qy 528 -----NSEDVVMHAIQLGNCVTITHTSRNDEFTPTSTVPVSFELNFYSNGVLHVFIM 582
Db 612 TVTPHTPERITIAAEQI--NILTRVHHPGLGDVLRVDGDNV-----LLSYFRNNVLHFTA 665
Qy 583 EAIACSLYAVLNKRGGLGGTSTPPNLIQSOBLVRKAASLCYLLSNGTISLPQCTFYQV 642
Db 666 SAWVACCFKN--NRR-----ISRLLIRLGVMPYFLOABLFLPWTEDQFAQH 711
Qy 643 CHETVGKFIQVGIITVAEHDDQEDISPSLAEOQWDKKLPEPLSWRSDEEDSDFGEEQR 702
Db 712 IQQVIELFVREGLL-LSAGDEED-----PLTRNTQOTDE----- 745
Qy 703 DCYLVKSQSKHEHQFITFLQRLGLLGLLEA-----YSSAAIFVHNFSGPVPEPEYLOKLHK 757
Db 746 -----VFLRAISHSLQAQAFERYITISILVKNPGGTLSASE-----LES 785
Qy 758 YLITRTERNVAVVAESATYCLVKNAVAMFKDIGVFKETKQK 798
Db 786 LCQLAAQRLSLLYASTA-----PEFFDKGLFRGFIQK 817

RESULT 7
US-10-369-493-9385
; Sequence 9385, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9385
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-369-493-9385

Query Match 11.7%; Score 500; DB 12; Length 754;
Best Local Similarity 22.1%; Pred. No. 1.3e-40;
Matches 194; Conservative 152; Mismatches 308; Indels 222; Gaps 26;

Qy 18 HSSEYSVGR-CKHTSEEWGEC-----GPRPTVFRSATLKWKESLMSRKPFGVGRCCYS 69
```



Db 3 HSYQVNDGRPIIVLEVDYGLCNTLILDKACRTKLPSPILPLDPNLPQKRAYLALSRSS 62  
Qy 70 CT-----PQSWDFFPNPSISL-----GLRNV-----IYINETHRHGHLA----- 106  
Db 63 SSNSLIPNQRGGKTHSDSLANLLOAHRIRDTLDVHLVPVSIPIGRTPDQSGWFAVLFS 122  
Qy 107 -----PQSWDFFPNPSISL-----GLRNV-----IYINETHRHGHLA----- 121  
Db 123 NWALVGRFRLLAVLNGRNTIVCFAPISVRQTLNEGLPPERTLRKLQVLETHFRIR 182  
Qy 122 KGMFATNVT-----ENVLNSSRVOEAEIAEVAELPDGSAQOQSKAVNKKAKRILO 175  
Db 183 ETWIGDLSFRLLVDNVLATEAVREAI---AAQAKEDGT-----DLSETWKAQAYAW 233  
Qy 176 EMVATVSPAMIRLTGWLKLFNSFFWNIHQLEEMVKAATETNLPPLFLPVHRSHID 235  
Db 234 EIAADYSSPVIRGADEFLFSVMNRIYAGVLIIH-HVDSFK-ETAPGHEVVYVPSHSHID 290  
Qy 236 YLLLTFLFCHNIKAPYIAGSNLNIPIFTLHKLGGFFIRRLDETDPGRKDVLYRAL 295  
Db 291 YMLLSYCLYCGGIVLPHIVAGINLNIPIVGTILRKCGAFFIRSI-----KGNMYSIV 344  
Qy 296 LHGHIVELLRQOQLEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPDILIIIPVGIS 355  
Db 345 LSEVVAQLVAGGYSLEYFIEGSRSTGRLLQPKGMMTTLQAFLRQPRPVLFPQIYIG 404  
Qy 356 YDRIIEG-HYNGEQLGPKKQNESLWSVARGVIRMLRKNYGCVRVDPAPPSLKKEYLESQS 414  
Db 405 YEKLIETSYLDELSGSPKESITWIRLFWNIPKLYKQYQVNVNFEPIALNDVL----- 460  
Qy 415 QKPVSALESLEQALLPAILRSPSDAEDGRTSINESRNATDESRLRRRLIANLAEHILF 474  
Db 461 -----AELAPEWEGQALNENEPALWS-----STVNHARQIQT 494  
Qy 475 TASKSCAIMSTHIVACLLYHRQGLDSTLVEDFFVMKEEVLARDPDLGSG----- 527  
Db 495 RINSADVNPINLALLALSTPKHAMGEADLIAQITLCKKILL-----ELFYSNRVTPH 550  
Qy 528 NSEDVVMHAIQLGNCVTIITHSRNDEFFITPTTVPSPFELNPNYNGVLHVFIMEAIIA 587  
Db 551 TPERITIAHAEQI--NILTRVHHPLGDLVLRVDGNAV-----LLSYFRNNVLHFTASAWA 604  
Qy 588 CSLVAVLNKRGIGGPTSTPNLISOEOLVRKAASCLVLLSNEGTSILPCQTFYQVCHETV 647  
Db 605 CCFKN--NRR-----MSRIALIRLGVMYFPLQAEFLPWTEDQFAHQHQQVI 650  
Qy 648 GKFTYQILTVABHDDQEDISPSLAEQOQWKLPPELSWSRSDDEDSDFGEQRDCYLK 707  
Db 651 ELFVREGLLSAGNEE-----DPLTRNTSQTDE----- 679  
Qy 708 VSOSKEHQOQITFLORLLGPLLEA-----YSSAAIFVHNFSGVPPEPYLQKLHKLITR 762  
Db 680 -----VFLRAISHLQOAFERYITISILVKNPGTILASE-----LESCLQA 724  
Qy 763 TERNVAVYASATVCLVKNVAKMFKDGVFKETKQK 798  
Db 725 AQLSLLYASTA-----PEFFDKLGRFGFIQK 751

## RESULT 8

US-10-369-493-13643  
; Sequence 13643, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 13643  
; LENGTH: 833  
; TYPE: PRT  
; ORGANISM: Pseudomonas fluorescens  
US-10-369-493-13643

Query Match 11.4%; Score 490; DB 12; Length 833;  
Best Local Similarity 22.4%; Pred. No. 1.5e-39;  
Matches 168; Conservative 149; Mismatches 264; Indels 170; Gaps 22;

Qy 88 LRNVYINETHRHGWLARRLSYVLFIOERDVHKGMPA-----TNVTENVLNSSRVOE 141  
Db 178 LRELEHNKGBHER---TVMAQRILRVHFRNLKAAVIGPDISHRRNLVKGLLNQPLVKQ 233  
Qy 142 AIAEVAALNPDGSAQOQSKAVNKKAKRILOEMVATVSPAMIRLTGWLKLFNSFF 201  
Db 234 AILD-----EAERENISEKAKAALRYGNEIASDYTYTAIR-----FLEVLSWF 279  
Qy 202 WNIQIHG---OLEMVKAATETNLPPLFLPVHRSHIDYLLTFLFCHNIKAPYIAGSN 257  
Db 280 WN-KIYDGIKVNHIEGVQKVAQGH-EVIYVPCRSHIDYLLSYLLFRNGLTPPHIAAGI 337  
Qy 258 NLNIPFSTLHKLGGFFIRRLDETDPGRKDVLYRALLHGHIVELLRQOQLEIFLEGT 317  
Db 338 NLNMPVIGLSLRGGAFMRFTKGNP-----LTSVFNELVHLLFTKGFVPEYVREG 391  
Qy 318 RRSRGKTSARAGLLSVVVDLTSTNVIPDILIIIPVGISYDRIIEGH-YNGEQLGPKKNE 376  
Db 392 RRTGEMLPQTKGMLAITLSRLSRMPVIVFPVYIGYVERVLEGTYLGELRGASKKE 451  
Qy 377 SLWSVARGVIRMLRKNYGCVRVDPAPPSLKKEYLESQSOKPVSALESLEQALLPAILPS- 435  
Db 452 SIFDIPK-VIGALKQRFQGVAVNFGPEPIKLAFLDSEQ-----PGW 491  
Qy 436 RPSDADEGRDTSINESRNATDESRLRRRLIANLAEHILFTASKSCAIMSTHIVACLLYR 495  
Db 492 ROELGFPQKPAWLNETTN-----RUGEKVAQHL-----NEAAAINPNVLVALLST 539  
Qy 496 HRQGLDSTLVEDFFVMKEEVLARDPDLGFS-----GNSDEVVMHA--IQ 538  
Db 540 TRALD-----DRAMARVLDLVALLRKVPYSPHTTLPEDGCRALIEHVKMD 587  
Qy 539 ILGNCVTIITHSRNDEFFITPTTVPSPFELNPNYNGVLHVFIMEAIIACSLYAVLNKRG 598  
Db 588 LL-----SEQNDALGKILYLDQNAVLMTYRNNVLHIFALPALLASFFQST----- 634  
Qy 599 LGGPTSTPNLISOEOLVRKAASCLVLLSNEGTSILPCQTFYQVCHETVGFQYQILTV 658  
Db 635 -----SRMSREQLRYTRALPYLOSELFIWTLDELDAVIDDOWLEAFVEQGLRAF 685  
Qy 659 AEHDDQEDISPSLAEQOQWKLPPELSWSRSDDEDSDFGEQRDCYLKVSOSKEHQOQFI 718  
Db 686 -----EKDYLAPSPSSRHVLL 703  
Qy 719 TFLQRLGLPRLLEYSSAAIFVHNFSGVPPEPYLQKLHKLITRTE-----RNVAVYAES 774  
Db 704 TLLSKSIAQTQRFYMTVSVLLNSGQNSISABELEDCTVMAQRILSHLHGLNAPPEFDKS 763  
Qy 775 TYCLVKNVAKMFKDGVFKETKQKRVSVLEL 805  
Db 764 ---LFRHF:QTLTLDLVLRRDEAGKLSYHEL 791

## RESULT 9

US-10-369-493-9240  
; Sequence 9240, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:



Db 604 LSVFRNVHLFTASSWVACCFQN--NRR-----MSRAGLRLGRVTVPFLQA 649

Qy 629 EGTISLPCOTFYQVCHETVKGKFIQYGILTVAEHDD 663

Db 650 ELFLPWSERFAERIEQTIDMFVREGLLLNVTDDD 684

RESULT 11

US-10-369-493-15639

; Sequence 15639, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 15639

; LENGTH: 792

; TYPE: PRT

; ORGANISM: Xanthomonas campestris

US-10-369-493-15639

Query Match 10.6%; Score 452; DB 12; Length 792;

Best Local Similarity 25.2%; Pred. No. 9.6e-36;

Matches 145; Conservative 113; Mismatches 223; Indels 94; Gaps 17;

Qy 107 RLRSYVLFTQERDVHKGFATNVT-----ENVLNSSRQVEAIAEVAELNPDGSAOQOS 160

Db 163 RKLQRLVTRHFRREAVIGPDLSTRLLVDQVLAADSVREAIAQAQRDN-----S 214

Qy 161 KAVNKYKKAKRLQEMVATVSPAMIRLTGWLLKLFNSFFWNIOIHKGQLEMYKAATET 220

Db 215 KPVD-AWRKAHAYAMEIAADYSPVRSASFLLTHVWNRIVAGVLVH--HLDKLUQAAPG 271

Qy 221 NLPLLELPVHRSHIDYLLMTFLFCHNIKAPYASGNLNIPFSTLHKLGQFFIRRL 280

Db 272 H-EVVYVPSHRSHMDYLLSYLLYERGVPHIVAGINLNPVGTLLRKGAFFIRSI 330

Qy 281 DETPDGRKQVLYRALHGHIVELLRQQOFLFLEGTSSRSKTSFCARAGLLSVVVDTL 340

Db 331 -----RGNALYSAVLSVYVAQVAGYSIEYFVEGGRSRTGRLLQPKGMIAMTLRAYL 384

Qy 341 TNVIPDILLIPVIGISYDRIIEGH-YNGEOLGPKKNESLWSVARGVIRMLRKNVGCVRVD 399

Db 385 RQRPKPVLPQVYVIGYKLEMEGNSYLDLTGRPKESIGWGLLWSIPKVLQKQYGVVVN 444

Qy 400 FAQPFSLKLEYLSQS-----QKPVSAALLSLEQALLPAILPSRPSDAADGRDTS 448

Db 445 FGEPIALNDVLAKHAEWNDEPLPDEKET-----WLAPAV-----DTLATOIQR 490

Qy 449 INESRNATDESURRLRIANLAHEILFTASKSCAIMSTHIVACLLYLRHQGIDILSTLVED 508

Db 491 INCAADVNPINLLAALLSTPKHAMEAD-----LIAQIELCKLLAEMPYSDRVTVT-- 543

Qy 509 FFWKVEVLARPDFLFGSGNSEDVWHAIQLLGNCVTITHTSRNDDEFITSTVPSVFE 568

Db 544 -----PHTPAR-----IITHABEI--NVLTRVSHPLGDLVLSGDTAV----L 580

Qy 569 LNFYSNGVLHVFTIMEAIIACSLYAVLNKRGGLGPTSTPPNLISOBOLVRKAASLYLLSN 628

Db 581 LSVFRNVHLFTASSWVACCFQN--NRR-----MSRAGLRLGRVTVPFLQA 626

Qy 629 EGTISLPCOTFYQVCHETVKGKFIQYGILTVAEHDD 663

Db 646 GAAEDRDKFVKVALETGRAEYHAGRITASESLAKVTILENAVAFLL 690

RESULT 14  
US-09-712-363-244  
; Sequence 244, Application US/09712363  
; Patent No. US20020164588A1  
; GENERAL INFORMATION:  
; APPLICANT: Eisenberg, David  
; APPLICANT: Rotstein, Sergio H.  
; APPLICANT: Marcotte, Edward M.  
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS  
; FILE REFERENCE: 07419-032001  
; CURRENT APPLICATION NUMBER: US/09/712,363  
; CURRENT FILING DATE: 2000-11-13  
; PRIOR APPLICATION NUMBER: PCT/US00/02246  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,531  
; PRIOR FILING DATE: 2000-02-01  
; PRIOR APPLICATION NUMBER: 60/117,844  
; PRIOR FILING DATE: 1999-01-29  
; PRIOR APPLICATION NUMBER: 60/118,206,  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: 60/126,593  
; PRIOR FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: 60/134,093  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/134,092  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/165,124  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/165,086  
; PRIOR FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 292  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 244  
; LENGTH: 789  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
US-09-712-363-244

Query Match 5.7%; Score 245; DB 10; Length 789;  
Best Local Similarity 22.0%; Pred. No. 7.4e-15;  
Matches 112; Conservative 95; Mismatches 206; Indels 96; Gap

QY 116 QERDV-----HKGMPATNTVENLSSRVQEAIAEAAE-----LNPDSGA  
DB 149 QWRDTTVAEHRKDFQAQFVSREAL-----LALAAEYRILGPQYKSPRLVVKPEMLA  
QY 159 QSKA-VNKVK---KKAKRIIQEUV---ATVSPAMIRLTGWVLLKLFNSFFWNIQIHKI  
DB 201 RFRAGLDRIPGATVEDAGKMLDELTGWSQVSDLVSVLGRLASRGDDEF---DYDE  
QY 211 LEMVKAATETNLPLLFLPVRSHDYLLLTILFELCHNIKAPYIASGNLNIPIFSTLII  
DB 258 VAAMEAALEAH-PAVLLFHSRHSYIDGVVVPVAMQDNRLPPVMEGGINLSFGLMGLPLM  
QY 271 LGGFPIRRRLDETPGRKDVLYRALLHGHIVELLRQQOFLFIIEGTRSRSGKTSKAR  
DB 317 SGMIPIRRNIGNDP-----LYKYVLKEYGVYVVEKFNLSWSIEGTRSRSTGRMLPPLK  
QY 331 LLSVVVDLTSTNVPDILLIPVGISYDRITIE-CHYNGEQKPKKKNESLWSVARGVIR  
DB 371 LMSYVADAYLGRSDPILLQGVSIICFDOLHEITEYAAAYARGAEKTPGLRLWLYNFKA  
QY 390 RKNYGVCRVDFAQPFSLKEYLESQKVPVSALLSLEQALLPAILPSRPSDAADEGRDT  
DB 431 ENFGKIVYRPEAVSMRQYL-----GAPHELTV  
QY 450 NESRNATDESURRLIANLAELHIFTASKSCAIMSTHIVACLLLYRHROGIDISTLVE  
DB 461 -----DPAKRIALQKMSFEVAWRILQATPVVATGIVSALLITTRGTALTIDOLHIA

QY 508 ----DFPMKEVLARDFDLGFGSNDSDVVMHAIQLL--GNCVTIHTSRNDEFFITPST 561  
Db 514 QDSLDYLERKQSPVST--ALRLRSRGVRAADALSNGHPVTRVDSGREPVWYIAPDD 570  
QY 562 TVPSVFELNFYSNGVLHVFIIMEAIIACSL 590  
Db 571 EHA AAA-----FYRNSVIAHFLTSIVELAL 595  
RESULT 15  
US-10-094-749-1905  
; Sequence 1905, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1905  
; LENGTH: 1095  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-094-749-1905

Query Match 3.0%; Score 128; DB 12; Length 1095;  
Best Local Similarity 19.2%; Pred.No. 0.0083;  
Matches 155; Conservative 118; Mismatches 329; Indels 204; Gaps 33;  
QY 114 FIOERDVHKGMAFNVTENVLNSRVOEATAEVAALNPDGSAQQSQKAVNKVKKAKRI 173  
Db 349 FLQNLVPHRSVSVTMMLEVVKNVSHVDVHTQGLVEL--GFILMDSYG-----PKKV 398  
QY 174 LOEWATVSPAMRLTGWLLK-----LFNSFFWNIQIHGQLEMY--KAATETNLPPLF 226  
Db 399 LDGKTETSPSLRMPNQHACKLANILLETFTKIHMIROEILEQVLNRVVTTRASSPI-- 456  
QY 227 LPVHRSHIDYLLTFLFCHNIKAPYIASGNLNI-----PI 263  
Db 457 -----SHFDLLSNIVMY-----APLVQSCSKVTEAFDYLSPLOTVQRLKAVQPL 506  
QY 264 FST-----LIHKLGGFFIRRLDETPDGRKDVLYREALHGHIVELLRQQQFLFLEG 316  
Db 507 LKVSMSMRDCLILVLRKAMFANQL-----DARKSAVAGFLL-----LLKNFKVL-----G 551  
QY 317 TRSRSGKTS CARGLLSVVVDTLSTNVIPDILIPGVSIDR-----IIEGHYN 365  
Db 552 SLSSQCSQSLSVSQVHVHSHYNSVANETFTCLEIMDSLRCLSSQADVRLMLYEGFYD 611

QY 366 GEQLGKPKKNESLW--SVARGVIRMLRKNGYCVRVDPAQPFSLKEYLESQS-----QKPV 419  
Db 612 -----VLRNSQLANSVMQTLLSQLKQFYE--PKPDLPLPKLEACTLTQGDKISLQEP 665  
QY 420 ALLSLEQALLP-----AILPSRPSDADEGRDTSINESRNATDESRLRLIANLAELHIFT 475  
Db 666 YLLCCIOHCLAWYKNTVIPLOQGESEEE--EEAFYEDLDLLESITNRMIKSELED--FE 722  
QY 476 ASKSCAINTSHIVACLLLYRHRQGDLSLTVLVEDFPVMK--BEVLARDFDL--CFSGNSE 532  
Db 723 LDKSAD-----FSQSTSIGIKNNISAFVLMGFCEVLIIEYNSISSFSKXRFED 770  
QY 533 VMHAIQLLGNVCVTIHTSRNDEFFITPSTTVPSVVFELNFYSNGVLHVFIIMEAIIACSL 592  
Db 771 ILSLFWCYKKLSDILNEKAGAKTKWANKTSDLSLSMKFVSS-----LLTALFR 819  
QY 593 VLNKRGLGGPTSTPPNLSIQEOLVRKAASLCLVLLSNEG--TISLPCQTFYQVCHETV 650  
Db 820 VLLWRYTSIPTSV-----EESGKKEGKGSISLLCLEGLQKIFSAVQOQFYQ---PKI 870  
QY 651 IQYGILTVAEHDDQEDISPSLAE-----QOWDKKLPEPLSWRSDEDESDSDFGEEQ 704  
Db 871 LRALDVTKEGEEREDADVSTORTAFQIRQFORSLLNLLS--SOREDFNS-----KEA 922  
QY 705 YLKVS-----QSKHQQFITTFLORLLGLPLLEYSSA 735  
Db 923 LLLVTVLTSLSKLLPSSPQFVQMLSWTSKICKENSREDALFCKSLMNLFLSLHVS 982  
QY 736 AIF-----VINFGVPPEPEYLOKLHKYLITRERNVAVVAESATYCLVKNVAMPKD 788  
Db 983 VILLRDLSDQIHGHLGDIDQDVEVEKTNHFAIV-----NLRTAAPTVCLLVLV 1038  
QY 789 IG-----VFKETKQKRVSL 805  
Db 1039 VDMLITKLKGQVSQETLSVSPGVSEL 1064

Search completed: January 10, 2004, 20:29:32  
Job time : 62 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2004, 19:36:42 ; Search time 70 Seconds  
(without alignments)  
3052.394 Million cell updates/sec

Title: US-09-935-290-2  
Perfect score: 4280  
Sequence: 1 MDESALTGTIDVSYLPHSS.....FLPQCNRKLLLEYLFSVVL 828

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	4265	99.6	828 4 Q8N1G6	Q8N1G6 homo sapien
2	3993.5	93.3	827 11 Q8VCT2	Q8VCT2 mus musculu
3	834	19.5	850 5 Q9Y137	Q9Y137 drosophila
4	831.5	19.4	757 5 Q8IMM7	Q8IMM7 drosophila
5	831.5	19.4	786 5 Q8IMM8	Q8IMM8 drosophila
6	442	10.3	809 16 Q8E809	Q8E809 shewanella
7	424	9.9	678 11 Q8R390	Q8R390 mus musculu
8	399.5	9.3	809 16 Q8DD48	Q8DD48 vibrio vuln
9	393	9.2	680 4 Q8BWC2	Q8BWC2 homo sapien
10	361	8.4	794 16 Q8D373	Q8D373 wiggleswort
11	353.5	8.3	663 5 Q9NHF2	Q9NHF2 drosophila
12	349.5	8.2	724 5 Q9VBQ6	Q9VBQ6 drosophila
13	337	7.9	671 5 Q9UZE8	Q9UZE8 caenorhabdi
14	335.5	7.8	320 11 Q8BRZ9	Q8BRZ9 mus musculu
15	178	4.2	809 16 Q8EYF2	Q8EYF2 leptospira
16	150	3.5	287 16 Q8YMR2	Q8YMR2 anabaena sp

17	140	3.3	3144	4	Q9UQB7	Q9ucp7 homo sapien
18	138.5	3.2	249	16	Q8PF86	Q8pf86 xanthomonas
19	135.5	3.2	239	16	Q97L74	Q97l74 clostridium
20	134	3.1	240	16	P73054	P73054 synechocyst
21	130.5	3.0	249	2	Q8EY25	Q8ey25 xanthomonas
22	129.5	3.0	249	16	Q8P3Q9	Q8p3q9 xanthomonas
23	129.5	3.0	1084	4	Q9NV11	Q9nv11 homo sapien
24	128	3.0	356	10	Q8GXU8	Q8gxu8 arabidopsis
25	122	2.9	654	4	Q86IV0	Q86iv0 homo sapien
26	120	2.8	1129	3	Q8J1G4	Q8j1g4 ashbya goss
27	118.5	2.8	906	4	Q96ST0	Q96st0 homo sapien
28	118	2.8	651	11	Q8K113	Q8k113 mus musculus
29	118	2.8	1258	4	Q00539	Q00539 homo sapien
30	118	2.8	1258	11	Q08982	Q08982 mus musculus
31	116	2.7	1262	3	Q13327	Q13327 candida alb
32	115.5	2.7	564	10	Q8L4Z8	Q8l4z8 oryza sativ
33	115	2.7	651	11	Q9J178	Q9j178 mus musculus
34	115	2.7	4133	13	Q9DE12	Q9dei2 gallus gall
35	115	2.7	4134	13	Q8QXK4	Q8gxk4 gallus gall
36	114	2.7	675	17	Q8TNP4	Q8tnp4 methanosarc
37	113	2.6	1231	4	Q8N3U4	Q8n3u4 homo sapien
38	112.5	2.6	1885	2	Q8RJY4	Q8rjy4 stigmatella
39	112	2.6	344	10	Q8W3E1	Q8w3e1 oryza sativ
40	112	2.6	822	5	P90531	P90531 dictyosteli
41	112	2.6	2273	3	Q42823	Q42823 saccharomyc
42	111.5	2.6	635	10	Q8S0J1	Q8s0j1 oryza sativ
43	111.5	2.6	805	5	Q95RK7	Q95rk7 drosophila
44	111.5	2.6	921	5	Q9VJ29	Q9vj29 drosophila
45	111	2.6	292	16	Q9KIH9	Q9kih9 rhizobium m

ALIGNMENTS

RESULT 1  
Q8N1G6 PRELIMINARY; PRT; 828 AA.  
ID Q8N1G6  
AC Q8N1G6;  
DT 01-OCT-2002 (TremBLrel. 22, Created)  
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)  
DE 01-MAR-2003 (TremBLrel. 23, Last annotation update)  
DE Hypothetical protein KIAA1560.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBDJ databases.  
DR EMBL; BC030783; AAH30783.1; -;  
DR InterPro; IPR002123; Acyltransferase.  
DR SMART; SM00563; PLSC; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 828 AA; 93874 MW; 55F6CF7EBCFE227A CRC64;

Query Match	99.6%	Score 4265;	DB 4;	Length 828;
Best Local Similarity	99.6%	Pred. No. 0;		
Matches 825;	Conservative 1;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1	MDESALTGTIDVSYLPHSSYSVGRCKHTSEEWGECGFRPTVFRSATLKWKESLMSRRK	60	
Db	1	MDEALTGTIDVSYLPHSSYSVGRCKHTSEEWGECGFRPTVFRSATLKWKESLMSRRK	60	
Qy	61	PFVGRCCVCTPQSWDKFFNPSPISLGLRNVIYINETHTRHGMRLARLSYVLFQERDV	120	
Db	61	PFVGRCCVCTPQSWDKFFNPSPISLGLRNVIYINETHTRHGMRLARLSYVLFQERDV	120	
Qy	121	HKGFMATNTVENLNSRRVQGAIAEVAELNPDGSAQQSKAVNKVKKKAKRILOEMVAT	180	
Db	121	HKGFMATNTVENLNSRRVQGAIAEVAELNPDGSAQQSKAVNKVKKKAKRILOEMVAT	180	

QY	181	VSPAMIRLTGWLLKLFNSFFWNIQIHKGQLEWKAATETNPLPLFLPVHRSHIDYLLLT	240
DB	181	VSPAMIRLTGWLLKLFNSFFWNIQIHKGQLEWKAATETNPLPLFLPVHRSHIDYLLLT	240
QY	241	FILFCHNIKAPYIAGNNLNIPISTLIHKLGGFFIRRRRLDETPDGRKDVLYRALLHGI	300
DB	241	FILFCHNIKAPYIAGNNLNIPISTLIHKLGGFFIRRRRLDETPDGRKDVLYRALLHGI	300
QY	301	VELLRQQOQFLEIFLEGTRSGKTSKARAGLLSVVVDTLSTNVIPIIDILIPVGSYDRII	360
DB	301	VELLRQQOQFLEIFLEGTRSGKTSKARAGLLSVVVDTLSTNVIPIIDILIPVGSYDRII	360
QY	361	EGHYNGQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPPSLKEYLESQKPVSA	420
DB	361	EGHYNGQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPPSLKEYLESQKPVSA	420
QY	421	LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC	480
DB	421	LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRRRLIANLAHILFTASKSC	480
QY	481	AIMSTHIVACLLYRHRQGDLSLTVDFVFMKEEVLARDFDLGFSGNSDVVWHAIQLL	540
DB	481	AIMSTHIVACLLYRHRQGDLSLTVDFVFMKEEVLARDFDLGFSGNSDVVWHAIQLL	540
QY	541	GNCVTITHTSRNDEFFITPTSTVPSVFNLFYNGVLHVFMIAIACSLYAVLNKRGGLG	600
DB	541	GNCVTITHTSRNDEFFITPTSTVPSVFNLFYNGVLHVFMIAIACSLYAVLNKRGGLG	600
QY	601	GPTSTPNLISQQLVKAASLCYLLSNEGTSISLPCQTFYQVCHETVKGFIQYGLTVAE	660
DB	601	GPTSTPNLISQQLVKAASLCYLLSNEGTSISLPCQTFYQVCHETVKGFIQYGLTVAE	660
QY	661	HDQOEDISPSLAEOQWDKKLPEPLSWRSDEDEDSDFGEQORDCYLKVQSKEHQOIFITF	720
DB	661	HDQOEDISPSLAEOQWDKKLPEPLSWRSDEDEDSDFGEQORDCYLKVQSKEHQOIFITF	720
QY	721	LQRLGLPALLEAYSSAAIFVHNFSGVPPEYVQKLVKYLITRTRNVAVVAESATYCLVK	780
DB	721	LQRLGLPALLEAYSSAAIFVHNFSGVPPEYVQKLVKYLITRTRNVAVVAESATYCLVK	780
QY	781	NAVMMFKDGVFKETQKRVSVLELSSTFLPQCNRQKLLLEYILSFVVL 828	
DB	781	NAVMMFKDGVFKETQKRVSVLELSSTFLPQCNRQKLLLEYILSFVVL 828	
RESULT 2			
ID	Q8VCT2	PRELIMINARY; PRT; 827 AA.	
AC	Q8VCT2;		
DT	01-MAR-2002 (TrEMBLrel. 20, Created)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Glycerol-3-phosphate acyltransferase, mitochondrial.		
GN	GPAM.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Kidney;		
RA	Strauberg R.;		
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC019201; AAH19201.1; -.		
DR	MGD; MGI:109162; Gpam.		
DR	InterPro; IPR002123; Acyltransferase.		
DR	SMART; SM00563; Plsc; 1.		
KW	Transferase; Acyltransferase.		
SQ	SEQUENCE 827 AA; 93704 MW; 4C177AA15374EE9B CRC64;		
Query Match 93.3%; Score 3993.5; DB 11; Length 827;			
Best Local Similarity 92.8%; Pred. No. 9e-317;			
Matches 768; Conservative 30; Mismatches 29; Indels 1; Gaps 1;			

```

RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier A., Fleischmann W.,
RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
RN [2]
RN SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcaina T., Baxter E., Blazej R.G., Butenhoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,
RA Celnik S.E.;
RT "Full length Drosophila melanogaster cDNA sequence."
RL Submitted (APR-1999) to the ENBL/GenBank/DBJ databases.
DR EMBL; AE003763; AA566761.1; -
DR EMBL; AF145635; AAD38610.1; -
DR FlyBase; FBgn0027579; BcDNA:GH07066.
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
DR SMART; SM00563; Pfam; 1.
SQ SEQUENCE 850 AA; 95164 MW; 54E569AB3B62A357 CRC64;

Query Match
Best Local Similarity 27.5%; Score 834; DB 5; Length 850;
Matches 230; Conservative 182; Mismatches 313; Indels 110; Gaps 22;

Qy 42 TVFSATLKWESLMSRRPFPVGRCCVCTPQSWDKFPNPSIPSLGRNVIYINETHR 101
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 45 SLFPNLNLVSYTSLSKINTIRSDLLTLRSSNEPKTONSAHGIGTQNMLEL---TPH 100
Qy 102 RGWLARLSY---VLFTQERDVHKG--MFATNVTENVLNSRVOEATAEVAE----- 149
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 101 SGQPFKFLFDWGLFPYVAQVVRSEKFEYRQVTEVHNDVLAHLKAIKQAEQTLREQRY 160
Qy 150 -----LNPDCSAQQQSK-----AVNKVKKKAKRIQEMVATVSPAMIRLT 189

```

## RESULT 4

```

Q81MM7 PRELIMINARY; PRT; 757 AA.
AC Q81MM7;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CG5508-08-PB.
GN BCDNA.GH07066.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

```





RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abrill J.F., Agbavani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Balaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Botkova D., Botchan M.R., Bouck J.J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glödek A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Labko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhou L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [2].  
 RP SEQUENCE FROM N.A.  
 RA Celisner S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferreria S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3].  
 RP SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celisner S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4].  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celisner S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5].  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003763; AAN14132.1;  
 SQ SEQUENCE 786 AA; 88312 MW; 6A3084B77993620 CRC64;

Query Match 19.4%; Score 831.5; DB 5; Length 786;  
 Best Local Similarity 28.9%; Pred. No. 9.1e-59;  
 Matches 215; Conservative 165; Mismatches 262; Indels 101; Gaps 19;  
 QY 129 VTEVNLSSRVOEIAEVAE-----LMPDGAQQQSK----- 161  
 DB 69 VTEVHNDVAVLKAHQAAEQTLREORYAKNGHRLLRSSASGDQEAQEEDEKRGISYQA 128  
 QY 162 AVNKKVKKAKKILCEWATVSPAMIRLTGWVLLKLFNSFFWNIIQIHKGQLEWMAATSTN 221  
 DB 129 ILRQEQRAISILKDMGSTLNGGLATSTWLYKLLPCFLSGVVTNTQIEMLKATATERS 188  
 QY 222 --LPLELPVHRSHIDYLLTFLIFCHNIKAPYIAGNNLNIPIFSLIHLKGFFFRRR 279  
 DB 189 PGTELIIVPLHRSHLDYIMVTWLTNNDIRSPLVAGNNLQIPVFGLLRGLGAFFIKRK 248  
 QY 280 LDETPDGRKOVLYRALLHGHIVELLRQOQFLEIFEGTSTRSGKTSACARAGLLSVVDTL 339  
 DB 249 IDPV-EGKQDLYRAALHLVTHALKQCHNVFFIEGGRTRTKGPKMPKGGILSVINAF 307  
 QY 340 STNVIPDILIPVGISYDRIIEGHYNGEOLGPKKNESLMSVARGVIRMLRKNYGCVRVD 399  
 DB 308 MDGSIIPALLVPVSVYERLVDFGNFVREQGEKKIPESFGKAIISGIWALKSNYGLMKRID 367  
 QY 400 FAQPFSLKYLESSQ-----KPVSAALLSLEQALLPAILSPRPSDADEGRDT 447  
 DB 368 FNEPYSIRELVNSYNKIAREGDNIAKYVKSARVLQHQ-----STSSLYGTDV 416  
 QY 448 SINESRNATDESRRRIIANAEHLITASKSCAIMSTHIVACLLLYVHRGID---LST 504  
 DB 417 VCEEHRN-----LIESISQVDFCAATSVMTNALAFLLTLFRNGABEQILSE 467  
 QY 505 LVEDFFVKSEVLARDFDLGSNSDVMHAIOLGNCVTIHTSRNDEFFITPSTVP 564  
 DB 468 ALDDL-----RNSLSCGCKDIGFSGESSQIVAYACDLGLVTRSRDENGRLIVKAVNSVE 523  
 QY 565 SVFELNFYSGVLHVFTIMEAIIACSLYAVL-----NKRGLGGPTSTPPNLSIQEQLVRKAA 620  
 DB 524 SFTELAYSNWLTPTHFALSSILLTTFPHSLPETENKEA-----VSRKLLIDTAL 574  
 QY 621 SLCYLLSNEGTSILPCTQTFVQVCHETVGKFIQYG-ILTVAEHDDQEDISPASLAQOQDKK 679  
 DB 575 ENCOIYRYEFTLNKPTQVLENLLYQQLDILLISGCVLTKLHDD---LPNGAE---GRR 627  
 QY 680 LPEPLSWRSEED-EDSDFGEEQORDCVLKVSQSEHQOQFITFLQRLGLPLEAYSSAA-- 736  
 DB 628 LANVLAECLEDDGVEDYRDGEADEPKLLFASETPSQORYIC---EVLAPFAMTYTVAQS 684  
 QY 737 -IFVHNFSGPVPEYLOKLHKYLIITERNVAVVAESATYCLVKNAVMPKDIQVFKET 795  
 DB 685 LQILHKNS--MLESEFISFVINDLSKVKRGSCIAESISTDSVRNCLLLEKKSIVIEVC 742  
 QY 796 KQKRVSVLEISSTFLPQCNRQKL 818  
 DB 743 NQOQMRLISLNTLY--EWSRESL 763  
 RESULT 6  
 Q8E8Q9  
 ID Q8E8Q9 PRELIMINARY; PRT; 809 AA.  
 AC Q8E8Q9;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 GN Glycyl-3-phosphate acyltransferase.  
 OS Shewanella oneidensis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
 OC Alteromonadaceae; Shewanella.  
 OX NCBI\_TaxID=70863;  
 RN [1]  
 RP SEQUENCE FROM N.A.



```

QY 766 NVAVVAESATYCLVKNVAKMFKDIGVFKETK 796
Db 613 GSSQCYDALSSSELQNALAAAFVRLGVVEKK 643

RESULT 8
Q8DD48
ID Q8DD48 PRELIMINARY; PRT; 809 AA.
AC Q8DD48;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glyceral-3-phosphate O-acyltransferase.
GN VV1165.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016800; AAC09636.1; -.
KW Acyltransferase; Transferase; Complete proteome.
SQ SEQUENCE 809 AA; 90596 MW; OE04FF0B4980DCDC CRC64;

Query Match 9.3%; Score 399.5; DB 16; Length 809;
Best Local Similarity 22.4%; Pred. No. 1.7e-23;
Matches 180; Conservative 145; Mismatches 223; Indels 257; Gaps 38;

QY 84 PSLGLRNVIYNEITH-----TEHGWLARRLSYLVFIQERDVHKG-----NPAIVNT 130
Db 169 PVWSLR-----YQAOSHGTSSIAHK-----LARVARHFSRQKLAAGDPDLPSSQVLF 218
QY 131 ENVLNRSRVQBAIAEVAALNPDGSAQOQSKAVNKKAKRILOEMVATVSPAMI----- 186
Db 219 -RLMKSPIAEOAIE-----EAKKNISMEKARKEADIMEIAADFSYLSVKGQD 268
QY 187 RLGTGWLKLFNSFFWNTQIHKGLEVMKAAETNLT-----PLFLPVHRSHIDYLLTF 241
Db 269 RLLGWL-----WN-KLYQG-LNINNAATVRLAQDGHGHEIVVPCVCHRSHMDYLLSY 317
QY 242 ILFCHNIKAPIASGNLNI-----PIFSTLIHKLGGFFIRRLDETPDGRKDVLYRALLH 297
Db 318 VLYHEGMVPPHTAAGINLNFPPAGPIF-----RRGGAFFIRRSF-----KGNRLYSTIFR 367
QY 298 GHIVELLRQOQFLEIFLEGTSRSRSGKTSARAGLLSVVVDLTSTNVIDPILIPVGISYD 357
Db 368 EYLAEFLPAKYSVEVFSGGSRSGRGLLPKATGMLAMTIQAWLGNRPVTLVPVYIGYE 427
QY 358 RIIE-GHYNGQLGPKPKNESLWSVARGVIRMLR--NYGCVRVDFDAPFSLKEYLESQS 414
Db 428 HVMVAVATYAKELGRKKEKEN-----AGLVRLTKLRNFGLYNGFGEPIPLNQYLYNEHA 483
QY 415 QK-----PVSALLSLEQALLPAILPSPSDAABEGRTSINESRNATDESRLRRLIANL 468
Db 484 PEWTKNDIDPMGA-----SRPQ-----WINP-----VVNQL 508
QY 469 AEHILFASKCAIMSTHIVACLLLYRHRQGDILSTLVED-----PFVMEKEVLARDPDLG 524
Db 509 ANKMWTHINDAANAULTCATAALASQRALSQKSLIHQIECYLQLLKNVPYKTYTPV 568
QY 525 FSGNSEVYVMAIQI-----LGNCVTITHSRNDEFITPTSTVPVSFELNFSYNG 575
Db 569 -SESAAEALVEHAISLDKFIETDWTGDIISL-----DRNQSLMT-----YRNN 612
QY 576 VLVHFMIEAITACSLYAVLNKRLGGTSTTPNIIISQOLVKAASLCYLLSNEGTSISLP 635
Db 613 IHLFALPSLIA-----QMIHQEN----- 632

```

```

QY 636 CQTFVQVCHETVGKFIQVGIILTVAEHDDQ-EDISPSLAEOQWDKKLPEPLSWRSDEDE- 693
Db 633 -----LTVSQIQQVAVIIPFLKAELEF-----LSHKEELDEL 665
QY 694 -----DSDFGEQRDCYLVKSQSKHEHQOQFITTFLQRLGLPLLEAYSSAAIFVHNFSG 744
Db 666 VVKVLNELVSQDLISLKED-----KVAQNQANTLTLLGRTTISSETLQRYSLA-----ENL 716
QY 745 PVPEPEYLQKLHKYLIITERNVAVAYAESATYCLVKNVAKMFKDIGVFKETKQKRVSVLE 804
Db 717 LVSNP-----LAKADLEQKSQDI-----AQRLRLHGINAPEYF-DKGVP-----AS 758
QY 805 LSSTFLPQ-----CNRQKLEY 821
Db 759 LFSTLKQOQYLDSDGNCDCGKTAQF 783

RESULT 9
Q9BWC2
ID Q9BWC2 PRELIMINARY; PRT; 680 AA.
AC Q9BWC2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glyceral-3-phosphate O-acyltransferase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC000450; AAH00450.1; -.
DR InterPro; IPR00123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
DR SMART; SM00563; Pfam; 1.
KW Acyltransferase; Transferase.
SQ SEQUENCE 680 AA; 77161 MW; A6EC9567D5693476 CRC64;

Query Match 9.2%; Score 393; DB 4; Length 680;
Best Local Similarity 25.6%; Pred. No. 4.4e-23;
Matches 141; Conservative 113; Mismatches 193; Indels 104; Gaps 24;

QY 68 YSCTPQSDKFPNPSIPSLGLRNVIYINETHTRHGWLARRLSYLVFIQERDVHKGMP-- 125
Db 25 YSKELKKWDEF-----EDIL-----EERRH-----VSDLKFMKCYTPLYKGITPC 66
QY 126 -ATNVTENVLNRSRVQBAIAEVAALNPDGSAQOQSKAVNKKAKRILOEMVATVSPA 184
Db 67 KPDIKCSVLNSEEIHVIKQLSKE-----SLOSVDVLRREEVSEILDEMSHKLRLG 117
QY 185 MIRLTGWLLKLFNSFFWNTQIHKGLEVMKAAETNLTNPLFLPVHRSHIDYLLTFL 244
Db 118 AIRFCATLTKSVFKQIFSKVCVNEEGIKLQALQEH-PVLLPSHRSYIDFLMLSLFLY 176
QY 245 CHNIKAPIASGNLNIPIFSTLIHKLGGFFIRRLDETPDGRKDVLYRALLHGHIVEL 303
Db 177 NYDLPVPVIAAGDFLGMKVMGELLRNMGGAFFMER-----TFGNGK--LYWAVFSEYVKTM 230
QY 304 LRQOQF-LEIFLEGTSRSRSGKTSARAGLLSVVVDLTSTNVIDPILIPVGISYDRIEG 362
Db 231 LRNGYAPVEFLEGTSRSRSGKTLTPKFGLLNIVMEPFKREVFDTYLVPISSYDKILEE 290
QY 363 H-YNGEOLGPKPKNESLWSVARGVIRMLRNKYNVCVRVDEFAQPFSLKEYLESQSOKPVSA 421
Db 291 TLVYVELLGVPKKESTTGLK-ARKILSNFSGSIHYFGDPVSLRSLAAGRMSRS----- 345
QY 422 LSLEQALLPAILPSPSDAABEGRTSINESRNA--TDESILRRL--IANLAHEILFTAS 477
Db 346 --SYNLVPRYIPQKQS-----EDMHAFVTEVAYKMLLOIENM----- 381

```



```

Db 389 -----ATGRLQFTPMAKQLAEQIDVRLKAHALCPQTMVAVPMLALQLYI 435
QY 641 QVCHETVGFIQVIGILTVAEHDDQEDISPSLAEQQWDKKLPEPLSWRSDEEDEDSDPGE 700
Db 436 NCFMFWLARPAPVALLAALKEQKQKQSTDIS---YDASL---CALHAHVTTMDALFOHE 489
QY 701 QRD CYLKVSQSKHEHQOFTIFLQRLGP-ILEAYSSAAIFVHN----- 741
Db 490 -----FIIESNREAAEFETHLQLLDERVVEVETSGRINVDNECSHVILAAAPFLCLY 544
QY 742 -----FSGP---VPPEYLOKL-----HKYLITRTRNVAVVAE 772
Db 545 YQLVTLRKPIPLEFESNKKELLVRQOQHVQQLQOPGASASHVHPYCLALDNLNIAIYAL 604
QY 773 SATYCLVKNVAMFKDQGVFKETQKQKVSLELSSTFLPQCNQKLEY 821
Db 605 IQRGYLVKS-----RDSGQMKIATPGKCLRELETQLEYCYQLMPFAQY 648

RESULT 12
Q9VBQ6 PRELIMINARY; PRT; 724 AA.
AC Q9VBQ6;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Glyceral-3-phosphate-acyltransferase protein.
GN DHAP-AT OR CG4625.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Fandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mohtrefi A.,
RA Mount S.M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,

```

RESULT 13

```

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
DR EMBL: AE003752; AAF56473.2; -.
DR FlyBase; FBgn0040212; Dhap-at-.
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
DR SMART; SM00563; Pfam; 1.
DR Acyltransferase; Transferase.
KW ACYLTRANSFERASE; TRANSFERASE.
SQ SEQUENCE 724 AA; 82524 MW; 5972629749CED804 CRC64;

Query Match 8.2%; Score 349.5; DB 5; Length 724;
Best Local Similarity 20.8%; Pred. No. 1.7e-19;
Matches 169; Conservative 135; Mismatches 279; Indels 228; Gaps 31;

QY 117 ERDVHKGMFATNTVENVL-----SSRVEATAEVAAE---LNP----- 152
Db 21 EQEOKTPTSAEAYWRNFKNITAPGNEASMTREFNPQVAISFEKYNLPKQLKHVLRSEKL 80
QY 153 ----DGSAAQQSKAVNKVKKAKRILOEMVATVSPAMIRLTGWVLL-----KLFNSFFWNI 204
Db 81 RSILEHYAKESGTPKQMERQARALIDIGLDRNWAIRWCGIAITAIGKRICDGFY--- 137
QY 205 QIHKGQLEMKVKAETNL-----PLLFLPVHRSHIDYLLTLFTLCHNTKAPYASGN 257
Db 138 -----VNSASMANVRKMGKCPVLYLPSPHSYMDFILMSYVICYYVDIEIPGIAAGM 188
QY 258 NLNTPI-PTSLIHLKLGFFIRRLDETPDGRKDVLYRALLHGHIVELLRQOQF-LEIFLE 315
Db 189 DFHSMFGMTWLRKTGAFMRERSFS-----NDELYMDIFREYMYALVANTHIGVEFFIE 242
QY 316 GTRSRSGKTS CARAGLLSVVVDTLSTNVI PDILLIPGISYDRIIEGH-YNGEOLGKPKK 374
Db 243 GTRSRNFKALVPKIGLLSMALLPYFTGEVDPNIVPVSVAVERVLEBQLFVYELGVPKP 302
QY 375 NESLWSVARGVIRMLR---KNYGCVRVDFAQPFSLKEYLESQSKPKVSALLSLEQALLPA 431
Db 303 KES-----TKGFFKALKIIDERFGKFLDFGEPISVKKEFF----- 337
QY 432 ILPSRPDAADGDRDTSINERNATDSLR-----RLIANLAHILFTASKSCAINTST 485
Db 338 -----GHDSAQRMQORAGVGGHLOKLNQREVELVKQLANEIYQQORRIVISTF 385
QY 486 HIVACLLYRHROGIDISTLVEDFVMKEEVLARDFDLFGSGNSEDVVMH---ATQLLGN 542
Db 386 NLLS-LYASQLVAQRSVLDE-----LARG-----VHLKRIFEQLG- 422
QY 543 CVTITHSRNDFFITPTSTVPSVFELNFGVNGVLHVIMEAIIACSLYAVLNKRGLOGP 602
Db 423 ---ARVS-----TTPSSIKADVIDAVEIHSNLFH-----ATGRL 454
QY 603 TSTP---PNLISQQLVR-KAASLCYLLSNGTISLPQCTFYQVCHETVGVKPIQVGIUTV 658
Db 455 QFTPMARQKLAEQIDVKRLKAHALCPQTMVAVPMLALQLYINCFMFWLARPVALLAAL 514
QY 659 AEHDDQEDISPSLAEQQWDKKLPEPLSWRSDEEDEDSDPGEQDCYLKVSQSKHEHQOFT 718
Db 515 KEQNNQKQSTDIS---YDASL---CALHAHVTTMDALFOHE-----FIIESNREAAEF 563
QY 719 TFLQRLGP-LLEAYSSAAIFVHN-----HKYLITRTRNVAVVAESATYCLVKNVAMFKD 744
Db 564 THLQQLLDERVVEVETSGRINVDNECSHVILAAAPFLCLYQVLTVLRKPIPLEFESN 623
QY 745 P---VPPEYLOKL-----HKYLITRTRNVAVVAESATYCLVKNVAMFKD 790
Db 624 KELLVRQOQHVQQLQOPGASASHVHPYCLALDNLNIAIYALIQRGYLVKS-----RDSG 678
QY 791 VFKEITKQKVSLELSSTFLPQCNQKLEY 821
Db 679 QMKIATPGKCLRELETQLEYCYQLMPFAQY 709

```







**THIS PAGE BLANK (USPTO)**

XX  
PT Kapeller-Libermann R. Logan T.J.

XX WPI; 2002-280934/32.  
 DR N-PSDB; ABK11094.  
 XX New acyltransferase-1, specifically glycerol-3-phosphate  
 PT acyltransferase, nucleic acids and proteins, useful for treating or  
 PT diagnosing e.g. metabolic disorders, weight disorders and appetite  
 PT regulation disorders  
 XX Claim 13; Fig 1; 122pp; English.  
 XX The invention relates to acyltransferase (ACTR-1) nucleic acids and  
 CC proteins. These are useful as modulating agents in regulating a variety  
 CC of cellular processes (e.g. cellular metabolism, energy homeostasis,  
 CC and/or cellular proliferation, growth, differentiation and/or migration).  
 CC The nucleic acids, proteins, protein homologues, fragments, antibodies,  
 CC peptides, peptidomimetics and small molecules can be used in screening  
 CC assays, predictive medicine (e.g. diagnostic or prognostic assays,  
 CC monitoring clinical trials, and pharmacogenetics), and in methods of  
 CC treatment (e.g. therapeutic or prophylactic). The ACTR-1 polypeptide or  
 CC its fragments are useful as reagents or targets in assays for the  
 CC treatment and/or diagnosis of ACTR-1 mediated or related disorders, which  
 CC include metabolic disorders such as disorders of energy homeostasis e.g.  
 CC diabetes, impaired glucose tolerance, insulin resistance, hyperglycaemia,  
 CC hypercholesterolaemia, hyperlipoproteinaemia, hypertriglyceridaemia,  
 CC and/or hyperlipidaemia; diabetic complications including atherosclerosis,  
 CC stroke, retinopathy, nephropathy, and peripheral neuropathy; weight  
 CC disorders and appetite regulation disorders, e.g. obesity, cachexia,  
 CC anorexia and bulimia. Proteins may also be used to screen for naturally  
 CC occurring ACTR-1 substrates, to screen for drugs or compounds which  
 CC modulate ACTR-1 expression and to treat disorders characterised by  
 CC insufficient or excessive production of ACTR-1 protein. The nucleic acids  
 CC may be used as hybridisation probes or primers, to express ACTR-1  
 CC protein, to detect ACTR-1 mRNA or a genetic alteration in an ACTR-1 gene,  
 CC to modulate ACTR-1 activity, in chromosome mapping, tissue typing and  
 CC forensic biology, and as surrogate markers (e.g. markers for precursors  
 CC of disease states, for predisposition of disease states, of drug  
 CC activity or of the pharmacogenomic profile of a subject). Modulators of  
 CC ACTR-1 protein activity or nucleic acid expression may be used to treat  
 CC a subject having a cardiovascular disorder or a triglyceride metabolism  
 CC disorder characterised by aberrant ACTR-1 protein activity or nucleic  
 CC acid expression. Anti-ACTR-1 antibodies can be used to detect and  
 CC isolate ACTR-1 proteins, regulate the bioavailability of ACTR-1  
 CC proteins, and modulate ACTR-1 activity. The present sequence  
 CC represents the amino acid sequence of human ACTR-1.  
 XX  
 SQ Sequence 828 AA;

Query Match 100.0%; Score 4280; DB 23; Length 828;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSALTLGTIDVSYLPHSSEYVGRCCKTSEWEGCGPPTVFRSATLKWESLSMRKR 60  
 DB 1 MDSALTLGTIDVSYLPHSSEYVGRCCKTSEWEGCGPPTVFRSATLKWESLSMRKR 60

QY 61 PFVGRCCYCTPQSDWKFFNPSPSLGLRNVIIYNETHRHGWLARRLSYVLFQERDV 120  
 DB 61 PFVGRCCYCTPQSDWKFFNPSPSLGLRNVIIYNETHRHGWLARRLSYVLFQERDV 120

QY 121 HKGMFATNTVENVLSNRVQEAIAEVAELNPDGSAQQQSKAVNKKVKKAKRIQEWAT 180  
 DB 121 HKGMFATNTVENVLSNRVQEAIAEVAELNPDGSAQQQSKAVNKKVKKAKRIQEWAT 180

QY 181 VSPAMIRLTGWLKLLFNSFFNIQIHKGQLEVMKAATETNPLFLPVHRSHIDYLLT 240  
 DB 181 VSPAMIRLTGWLKLLFNSFFNIQIHKGQLEVMKAATETNPLFLPVHRSHIDYLLT 240

QY 241 FILFCHNIKAPYASGNLNIPIFSTLIHKLGGFFIRRLDETPDGKQVLYRALLHGI 300  
 DB 241 FILFCHNIKAPYASGNLNIPIFSTLIHKLGGFFIRRLDETPDGKQVLYRALLHGI 300

QY 301 VELLRQOQFLEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPDILIPVGISYDRII 360

DB 301 VELLRQOQFLEIFLEGTRSRSGKTSARAGLLSVVVDLTSTNVIPDILIPVGISYDRII 360  
 QY 361 EGHYNGEQLGKPKNESLWSVARGVIRMLRKNYGCVRVDFQAFPSLKEYLESOSQKPVSA 420  
 DB 361 EGHYNGEQLGKPKNESLWSVARGVIRMLRKNYGCVRVDFQAFPSLKEYLESOSQKPVSA 420  
 QY 421 LLSLQALLPALPSRPSDAADDEGRDTSINESRNATDESRLRLIANLAELHILFTASKSC 480  
 DB 421 LLSLQALLPALPSRPSDAADDEGRDTSINESRNATDESRLRLIANLAELHILFTASKSC 480  
 QY 481 AIMSTHIVACLLLYRHRQIGIDLTSLVEDFFVMKEVLARDPDLGFSGNSDEVVMAIQLL 540  
 DB 481 AIMSTHIVACLLLYRHRQIGIDLTSLVEDFFVMKEVLARDPDLGFSGNSDEVVMAIQLL 540  
 QY 541 GNCVTITHTSRNDEFFITPSTTPSVFELNFSVGLVHVFIMEAIIACSLYAVLNKRGIG 600  
 DB 541 GNCVTITHTSRNDEFFITPSTTPSVFELNFSVGLVHVFIMEAIIACSLYAVLNKRGIG 600  
 QY 601 GFTSTPPNLIISOQLVRKAASLCYLLSNEGTTSLPCQTFYQVCHETVQKFIQVGLITVAE 660  
 DB 601 GFTSTPPNLIISOQLVRKAASLCYLLSNEGTTSLPCQTFYQVCHETVQKFIQVGLITVAE 660  
 QY 661 HDDQEDISPSLAEQWQDKKLPEPLSWRSDEDEDESDFGEEQDCYLKVSQSKEHQOFTIF 720  
 DB 661 HDDQEDISPSLAEQWQDKKLPEPLSWRSDEDEDESDFGEEQDCYLKVSQSKEHQOFTIF 720  
 QY 721 LQRLGLPALLEAYSSAAIFVHNFSGVPPEPYLOKLHKYLITERNVAVYASATYCLVK 780  
 DB 721 LQRLGLPALLEAYSSAAIFVHNFSGVPPEPYLOKLHKYLITERNVAVYASATYCLVK 780  
 QY 781 NAVKFKDIDGVKTKOKRVSVLELSTPLPCNQKQKLELYLSFWVL 828  
 DB 781 NAVKFKDIDGVKTKOKRVSVLELSTPLPCNQKQKLELYLSFWVL 828

## RESULT 2

ABG66665  
 ID ABG66665 standard; Protein; 828 AA.

AC ABG66665;

DT 29-AUG-2002 (first entry)

DE Human glycerol-3-phosphate acyltransferase hGPAT.

KW Fatty acid regulated gene; polyunsaturated fatty acid disorder;  
 KW PUFA disorder; eczema; cardiovascular disorder; hypertriglyceridaemia;  
 KW dyslipidaemia; atherosclerosis; coronary artery disease;  
 KW cerebrovascular disease; peripheral vascular disease; inflammation;  
 KW sinusitis; asthma; pancreatitis; osteoarthritis; rheumatoid arthritis;  
 KW acne; body weight disorder; obesity; cachexia; anorexia;  
 KW psychiatric disorder; cancer; cystic fibrosis; pre-menstrual syndrome;  
 KW diabetic; diabetic complication; genetic polymorphism.

OS Homo sapiens.

XX WO200240666-A2.

XX 23-MAY-2002.

XX 19-NOV-2001; 2001WO-CA01632.

XX 17-NOV-2000; 2000US-248589P.

XX (XENO-) XENON GENETICS INC.

PI Winther MD, Goldberg YP, Knickle LC, Haardt M, Allen SJ, Ponton A;  
 PI De Antueno RJ, Jenkins DK, Nwaka SO;

XX WPI; 2002-508327/54.  
 DR N-PSDB; ABK94821.

PT Novel isolated polypeptide segment encoded by fat regulated genes,  
PT useful for diagnosing the presence of or a predisposition for a  
PT disorder involving fatty acid regulated genes in a subject  
PS  
XX Claim 14; Fig 28; 225pp; English.  
CC The invention describes an isolated polypeptide segment (I) whose genes  
CC are fat regulated. (I) or the polynucleotide encoding it (II) are useful  
CC for diagnosing the presence of or a predisposition for a disorder  
CC involving fatty acid regulated genes in a subject. A composition  
CC containing (I) or (II) is useful for treating a disorder involving fatty  
CC acid regulated genes, where the disorder is selected from a  
CC polyunsaturated fatty acid (PUFA) disorder, eczema, cardiovascular  
CC disorders (such as hypertriglyceridaemia, dyslipidaemia, atherosclerosis,  
CC coronary artery disease, cerebrovascular disease or peripheral vascular  
CC disease), inflammation (such as sinusitis, asthma, pancreatitis,  
CC osteoarthritis, rheumatoid arthritis or acne), body weight disorders  
CC (such as obesity, cachexia or anorexia), psychiatric disorders, cancer,  
CC cystic fibrosis, pre-menstrual syndrome, diabetes, and diabetic  
CC complications. (I) or (II) is useful as research agent and materials for  
CC discovery of treatments and diagnostics for a disease, particularly human  
CC disease. (II) is useful for constructing nucleotide probes and primers,  
CC for detecting genetic polymorphism, for detecting changes in the level of  
CC expression of (II), and as a diagnostic tool. This is the amino acid  
CC sequence of a protein regulated by fatty acids.

SQ Sequence 828 AA;

Query Match 100.0%; Score 4279; DB 23; Length 828;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 827; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDESALTGTIDVSYLPHSSYSVGRCKHTSEEWGECGFRPTFRSATLKWKESLMGRKR 60  
DB 1 MDESALTGTIDVSYLPHSSYSVGRCKHTSEEWGECGFRPTFRSATLKWKESLMGRKR 60  
QY 61 PFVGRCCYCTPQSWDKFNFPSLGLRNVIYNETHRHGWLARLSVYLFQERDV 120  
DB 61 PFVGRCCYCTPQSWDKFNFPSLGLRNVIYNETHRHGWLARLSVYLFQERDV 120  
QY 121 HKGFATNTENVLNSSRVQSAIEVAELNPDGSAQOOSKAVNKVKKAKRIILOEMVAT 180  
DB 121 HKGFATNTENVLNSSRVQSAIEVAELNPDGSAQOOSKAVNKVKKAKRIILOEMVAT 180  
QY 181 VSPAMIRLTGWLLKLFNSFFWNIQHKQLEWKAATETNLPILFLPVHRSHIDYLLLT 240  
DB 181 VSPAMIRLTGWLLKLFNSFFWNIQHKQLEWKAATETNLPILFLPVHRSHIDYLLLT 240  
QY 241 FILFCHNIKAPYIASGNNLNIPFSTLIHKLGGFFIRRLDETDPGRKDVLYRALLHGH 300  
DB 241 FILFCHNIKAPYIASGNNLNIPFSTLIHKLGGFFIRRLDETDPGRKDVLYRALLHGH 300  
QY 301 VELLRQOQFLFLEGTSSRGKTSARAGLLSVVVDLTSTNVPDILIIIPVGSYDRII 360  
DB 301 VELLRQOQFLFLEGTSSRGKTSARAGLLSVVVDLTSTNVPDILIIIPVGSYDRII 360  
QY 361 EGHYNGEQLGPKKKNESLWSVARGVIRMLRKNYGCVRVDFAPPSLKEYLESQKPVSA 420  
DB 361 EGHYNGEQLGPKKKNESLWSVARGVIRMLRKNYGCVRVDFAPPSLKEYLESQKPVSA 420  
QY 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSC 480  
DB 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSC 480  
QY 481 AIMSTHIVACLLLYHRRGIDLTIVDFFWKKEVLARDDFLGSNGSDVVMHAIQLL 540  
DB 481 AIMSTHIVACLLLYHRRGIDLTIVDFFWKKEVLARDDFLGSNGSDVVMHAIQLL 540  
QY 541 GNCVTITHTSRNDEFTPTVPSEFNFVSNGLVHVFIMEALIASCLYAVLNKGLG 600  
DB 541 GNCVTITHTSRNDEFTPTVPSEFNFVSNGLVHVFIMEALIASCLYAVLNKGLG 600  
QY 601 GPTSTPPNLIHQEQVLVKAASLCYLLSNEGTSISLPQCTFYQVCHETVKGFIQYGLITVAE 660

Db 601 GPTSTPPNLIHQEQVLVKAASLCYLLSNEGTSISLPQCTFYQVCHETVKGFIQYGLITVAE 660  
QY 661 HDQEDISPSIAEQWMDKLPPLSWRSDEDESDSDFGEQRDCYLKVSQSKHQQTTF 720  
Db 661 HDQEDISPSIAEQWMDKLPPLSWRSDEDESDSDFGEQRDCYLKVSQSKHQQTTF 720  
QY 721 LQRLGLPLLAYSAALFVNFSGPVPPEPYLQKLHKYLITRTRNNVAVVAESATYCLVK 780  
Db 721 LQRLGLPLLAYSAALFVNFSGPVPPEPYLQKLHKYLITRTRNNVAVVAESATYCLVK 780  
QY 781 NAVKMFIDIGVFKETKOKRVSVLELSSTFLPQCNROKLELYILSFVVL 828  
Db 781 NAVKMFIDIGVFKETKOKRVSVLELSSTFLPQCNROKLELYILSFVVL 828

RESULT 3

AAE22144  
ID AAE22144 standard; Protein; 828 AA.

XX  
AC AAE22144;

XX  
DT 25-JUL-2002 (first entry)

XX  
DE Human TRNFR-6 protein.

XX  
KW Human; transferase; developmental disorder; cell proliferative disorder;  
KW TRNFR-6; neurological disorder; autoimmune disorder; parasitic infection;  
KW inflammatory disorder; endocrine; antiparasitic; immunosuppressive;  
KW cytostatic; neurological.

OS  
FH Homo sapiens.

FT  
FT Domain 172..197 Location/Qualifiers  
FT Domain 215..412 /note= "Transmembrane domain"  
FT /note= "Acetyltransferase domain"

PN  
PD WO200226950-A2.

PF  
PF 04-APR-2002.

PF  
PF 28-SEP-2001; 2001WO-US30424.

PF  
PF 29-SEP-2000; 2000US-236523P.

PF  
PF 06-OCT-2000; 2000US-238481P.

PF  
PF 27-OCT-2000; 2000US-244025P.

PF  
PF 03-NOV-2000; 2000US-246001P.

PF  
PF 09-NOV-2000; 2000US-247931P.

PF  
PF 16-NOV-2000; 2000US-249639P.

PF  
PF 21-NOV-2000; 2000US-252819P.

PF  
PF (INCY-) INCYTE GENOMICS INC.

PI  
PI Lal PG, Tang YT, Yue H, Burford N, Gandhi AR, Warren BA, Yao MG;  
PI Tribouley CM, Baughn MR, Lee EA, Hafalia AJA, Lu Y, Griffin JA;  
PI Sanjanwala MS, Ding L,

PI  
PI WPI; 2002-362492/39.

PI  
PI N-PSDB; AAD35221.

PI  
PI Novel human transferase polypeptides and polynucleotides, useful in  
PI treating e.g., cell proliferative and autoimmune disorders -  
PI Claim 61; Page 126-128; 168pp; English.

CC  
CC The present invention relates to novel human transferases (TRNFR) and  
CC polynucleotides encoding such proteins. The TRNFR proteins are useful  
CC for treating disorders associated with a decreased expression of  
CC functional TRNFR, e.g., cell proliferative, developmental, neurological,  
CC autoimmune/inflammatory disorders and parasitic infections. Antagonists  
CC of TRNFR proteins are useful in treating disorders associated with

CC increased activity of TRNFR. The present sequence is human TRNFR-6  
 CC protein.  
 XX  
 SQ Sequence 828 AA;  
 Query Match 100.0%; Score 4279; DB 23; Length 828;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 827; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDESALTGLTIDVSYLPHSSEYSGRCCKHTSEWGECEGFPRTVFRSATLKKESLSMRKR 60  
 DB 1 MDESALTGLTIDVSYLPHSSEYSGRCCKHTSEWGECEGFPRTVFRSATLKKESLSMRKR 60  
 QY 61 PFVGRCCYCTPQSDWKFFNPSPISGLRNVIYINETHTRHGWLAARLSYVLFIQERDV 120  
 DB 61 PFVGRCCYCTPQSDWKFFNPSPISGLRNVIYINETHTRHGWLAARLSYVLFIQERDV 120  
 QY 121 HKGMFATNTENVLNSRRVQEAIAEVAALNPDGSAQQQSKAVNKKAKRILQEMVAT 180  
 DB 121 HKGMFATNTENVLNSRRVQEAIAEVAALNPDGSAQQQSKAVNKKAKRILQEMVAT 180  
 QY 181 VSPAMIRLTGWLLKLFNSFFWNIQHKQLEWVKAATETNLPFLPVHRSHIDYLLLT 240  
 DB 181 VSPAMIRLTGWLLKLFNSFFWNIQHKQLEWVKAATETNLPFLPVHRSHIDYLLLT 240  
 QY 241 FILFCHNIKAPYIAGNNLNIPFSTLIHKLGFFIRRRDETPDGRKOVLYRALLHGI 300  
 DB 241 FILFCHNIKAPYIAGNNLNIPFSTLIHKLGFFIRRRDETPDGRKOVLYRALLHGI 300  
 QY 301 VELLROOQFLEIFLEGTRSRSGKTSARAGLLSVVDTLSTNVIPIIDILIPVGISYDRII 360  
 DB 301 VELLROOQFLEIFLEGTRSRSGKTSARAGLLSVVDTLSTNVIPIIDILIPVGISYDRII 360  
 QY 361 EGHNGEOLGPKPKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQSPVSA 420  
 DB 361 EGHNGEOLGPKPKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQSPVSA 420  
 QY 421 LLSLEQALLPAILPSRPSDAADGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
 DB 421 LLSLEQALLPAILPSRPSDAADGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
 QY 481 AIMSTHIVACLLYRHRQGDISTLVEDFFVMKEEVLARDFDLFGSGNSEDVVMHAIQLL 540  
 DB 481 AIMSTHIVACLLYRHRQGDISTLVEDFFVMKEEVLARDFDLFGSGNSEDVVMHAIQLL 540  
 QY 541 GNCVTITHTSRNDEFFITPTSTVPSPELNFYNGVLHVFMIAIACSLYAVLNKRGILG 600  
 DB 541 GNCVTITHTSRNDEFFITPTSTVPSPELNFYNGVLHVFMIAIACSLYAVLNKRGILG 600  
 QY 601 GPTSTPNLISQOLVRKAASLCYLLSNEGTSIPLCQTFYQVCHETVKGFIQYGLITVAE 660  
 DB 601 GPTSTPNLISQOLVRKAASLCYLLSNEGTSIPLCQTFYQVCHETVKGFIQYGLITVAE 660  
 QY 661 HDQEDISPSLAQOMDKLPELSWRSDEEDSDSFGGEQRDCYLKVSQKEHQFIF 720  
 DB 661 HDQEDISPSLAQOMDKLPELSWRSDEEDSDSFGGEQRDCYLKVSQKEHQFIF 720  
 QY 721 LQRLGLPLLEYSSAAIFVNFSGPPEPEYLOKLHKYLIITRERNVAVAESATYCLVK 780  
 DB 721 LQRLGLPLLEYSSAAIFVNFSGPPEPEYLOKLHKYLIITRERNVAVAESATYCLVK 780  
 QY 781 NAVMFKDIGVFKETQKRVSVLELSSTFLPQCNROKLEIYILSFVVL 828  
 DB 781 NAVMFKDIGVFKETQKRVSVLELSSTFLPQCNROKLEIYILSFVVL 828

RESULT 4  
 AA027073  
 ID AA027073 standard; Protein; 828 AA.  
 XX  
 AC AA027073;  
 DT 22-MAY-2003 (first entry)

XX GPAM related protein sequence, SEQ ID No 2.  
 DE Antidiabetic; nephrotropic; neuroprotective; ophthalmological; human;  
 KW mitochondrial sn-glycerol-3-phosphate acyltransferase; GPAM;  
 KW diabetic complication; retinopathy; neuropathy; enzyme.  
 XX Homo sapiens.  
 XX WO2003008590-A1.  
 XX 30-JAN-2003.  
 XX 16-JUL-2002; 2002WO-JP071189.  
 XX 16-JUL-2001; 2001JP-0215337.  
 XX (KISP) KISSEI PHARM CO LTD.  
 XX Sakamoto S, Onota H, Sugano S, Nakamura Y;  
 XX WPI; 2003-229583/22.  
 DR N-PSDB; AAL55475.  
 XX Human mitochondrial sn-glycerol-3-phosphate acyltransferase and  
 PT antagonists for treatment and prevention of diabetic complications -  
 XX Disclosure; Page 40-46; 56pp; Japanese.  
 CC The invention relates to a novel protein having human mitochondrial sn-  
 CC glycerol-3-phosphate acyltransferase (GPAM) activity. The novel protein  
 CC with GPAM activity can be used in the prevention and treatment of  
 CC diabetic complications, including retinopathy and neuropathy, by  
 CC administration of antagonists to human GPAM. This sequence represents a  
 CC human protein relating to the GPAM activity protein of the invention.  
 XX Sequence 828 AA;

Query Match 99.9%; Score 4275; DB 24; Length 828;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 827; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MDESALTGLTIDVSYLPHSSEYSGRCCKHTSEWGECEGFPRTVFRSATLKKESLSMRKR 60  
 DB 1 MDESALTGLTIDVSYLPHSSEYSGRCCKHTSEWGECEGFPRTVFRSATLKKESLSMRKR 60  
 QY 61 PFVGRCCYCTPQSDWKFFNPSPISGLRNVIYINETHTRHGWLAARLSYVLFIQERDV 120  
 DB 61 PFVGRCCYCTPQSDWKFFNPSPISGLRNVIYINETHTRHGWLAARLSYVLFIQERDV 120  
 QY 121 HKGMFATNTENVLNSRRVQEAIAEVAALNPDGSAQQQSKAVNKKAKRILQEMVAT 180  
 DB 121 HKGMFATNTENVLNSRRVQEAIAEVAALNPDGSAQQQSKAVNKKAKRILQEMVAT 180  
 QY 181 VSPAMIRLTGWLLKLFNSFFWNIQHKQLEWVKAATETNLPFLPVHRSHIDYLLLT 240  
 DB 181 VSPAMIRLTGWLLKLFNSFFWNIQHKQLEWVKAATETNLPFLPVHRSHIDYLLLT 240  
 QY 241 FILFCHNIKAPYIAGNNLNIPFSTLIHKLGFFIRRRDETPDGRKOVLYRALLHGI 300  
 DB 241 FILFCHNIKAPYIAGNNLNIPFSTLIHKLGFFIRRRDETPDGRKOVLYRALLHGI 300  
 QY 301 VELLROOQFLEIFLEGTRSRSGKTSARAGLLSVVDTLSTNVIPIIDILIPVGISYDRII 360  
 DB 301 VELLROOQFLEIFLEGTRSRSGKTSARAGLLSVVDTLSTNVIPIIDILIPVGISYDRII 360  
 QY 361 EGHNGEOLGPKPKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQSPVSA 420  
 DB 361 EGHNGEOLGPKPKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQSPVSA 420  
 QY 421 LLSLEQALLPAILPSRPSDAADGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480  
 DB 421 LLSLEQALLPAILPSRPSDAADGRDTSINESRNATDESRRRLIANLAHILFTASKSC 480

QY 481 AIMSTHIVACLLLYRHROGIDILSTLVEDFFVMKEEVLARDFDLGFSGNSEDDVWHAIQLL 540  
DB 481 AIMSTHIVACLLLYRHROGIDILSTLVEDFFVMKEEVLARDFDLGFSGNSEDDVWHAIQLL 540  
QY 541 GNCVTITHTSRNDEFFITPSTTPVPSVFEINLFYSNGVLHVFMIALIACSLYAVLNKRGGLG 600  
DB 541 GNCVTITHTSRNDEFFITPSTTPVPSVFEINLFYSNGVLHVFMIALIACSLYAVLNKRGGLG 600  
QY 601 GPTSTPPNLIISOQLVRKAASLCYLLSNEGTSILPCQTFYQVCHETVKGFIQYGLITVAE 660  
DB 601 GPTSTPPNLIISOQLVRKAASLCYLLSNEGTSILPCQTFYQVCHETVKGFIQYGLITVAE 660  
QY 661 HDDQEDISPSLAEOQWKKLPEPLSWRDEDEDDSDFGEEQDCYLVKVSQKHEHQOFTTF 720  
DB 661 HDDQEDISPSLAEOQWKKLPEPLSWRDEDEDDSDFGEEQDCYLVKVSQKHEHQOFTTF 720  
QY 721 LQRLIGPLLEAYSSAAIFVHNFSGVPPEYLOKLHKYLIITRTENNAVYAESATYCLV 780  
DB 721 LQRLIGPLLEAYSSAAIFVHNFSGVPPEYLOKLHKYLIITRTENNAVYAESATYCLV 780  
QY 781 NAVRMFKDYGFKETKQKRVSVLELSSTFLPQCNQKLLVILSPVVL 828  
DB 781 NAVRMFKDYGFKETKQKRVSVLELSSTFLPQCNQKLLVILSPVVL 828

RESULT 5

ID ABUS4605  
XX ABUS4605 standard; Protein; 828 AA.

AC ABUS4605;

DT 03-JUN-2003 (first entry)

DE Human NOVX polypeptide #64.

XX Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;  
KW hypertension; congenital heart defect; aortic stenosis; valve disease;  
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;  
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;  
KW tuberosus sclerosis; scleroderma; atherosclerosis; infectious disease;  
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;  
KW Parkinson's disease; immune disorder; haematopoietic disorder;  
KW haemophilia; hypercoagulation; Crohn's disease; cancer.

OS Homo sapiens.

PN WO200281498-A2.

XX 17-OCT-2002

XX 03-APR-2002; 2002WO-US10780.

XX 03-APR-2001; 2001US-281086P.

XX 05-APR-2001; 2001US-281136P.

XX 05-APR-2001; 2001US-281863P.

XX 06-APR-2001; 2001US-281906P.

XX 10-APR-2001; 2001US-282020P.

XX 10-APR-2001; 2001US-282930P.

XX 12-APR-2001; 2001US-282934P.

XX 13-APR-2001; 2001US-283710P.

XX 17-APR-2001; 2001US-284234P.

XX 19-APR-2001; 2001US-285325P.

XX 20-APR-2001; 2001US-285381P.

XX 20-APR-2001; 2001US-285609P.

XX 23-APR-2001; 2001US-285748P.

PR 30-MAY-2001; 2001US-294484P.  
PR 18-JUN-2001; 2001US-298952P.  
PR 19-JUN-2001; 2001US-299237P.  
PR 19-JUN-2001; 2001US-299276P.  
PR 12-SEP-2001; 2001US-318750P.  
PR 25-SEP-2001; 2001US-324800P.  
PR 25-SEP-2001; 2001US-324802P.  
PR 27-SEP-2001; 2001US-325684P.  
PR 17-OCT-2001; 2001US-330143P.  
PR 14-NOV-2001; 2001US-332131P.  
PR 14-NOV-2001; 2001US-332240P.  
PR 14-NOV-2001; 2001US-332779P.  
PR 21-NOV-2001; 2001US-332115P.  
PR 04-DEC-2001; 2001US-337621P.  
PR 03-JAN-2002; 2002US-345783P.  
PR 16-JAN-2002; 2002US-350251P.  
PR 02-APR-2002; 2002US-0114270.

(CURA-) CURAGEN CORP.

Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA, Patturajan M;  
Liu X, Gusev VY, Li L, Vernet CAM, Zerhusen BD, Gorman L;  
Shenoy SG, Pena CEA, Smithson G, Burgess CE, Gerlach V;  
Padigaru M, Shimkets RA, Gangolli EA, Taupier RJ, Casman SJ, Ji W;  
Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;  
MacDougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;  
Ellerman K;

WPI; 2003-046858/04.  
N-PSDB; ABX72233.

New isolated NOVX polypeptide useful for treating atherosclerosis,  
metabolic disorders, diabetes, obesity, infectious disease, anorexia,  
neurodegenerative disorders, Alzheimer's disease and cancer

Claim 1; Page 232; 666pp; English.

The invention relates to human polypeptides, termed NOVX, and the  
polynucleotides encoding them. The polypeptides and polynucleotides are  
useful for diagnosing disease, and screening for potential therapeutic  
agents. The sequences are useful for treating metabolic disorders,  
cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic  
stenosis, atrial septal defect (ASD), atrioventricular canal defect,  
ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular  
septal defect (VSD), valve diseases, tuberosus sclerosis, scleroderma,  
atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative  
disorders, Alzheimer's disease, Parkinson's disease, immune disorders,  
haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease  
and cancer. Sequences ABUS4542-ABUS4647 represent human NOVX polypeptides  
of the invention.

Sequence 828 AA;

Query Match 99.8%; Score 4270; DB 24; Length 828;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 826; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDESALTLCITDVSYPHSSEYSVGRCKHTSEWEGCGFRPTVPRSATLWKESLMSRKR 60

Db 1 MDESALTLCITDVSYPHSSEYSVGRCKHTSEWEGCGFRPTVPRSATLWKESLMSRKR 60

Qy 61 PFVGRCCYCTPOSWDKFFNPSPSLGLRNVIIYNETHTRHGLARLSVYLFQIQRDV 120

Db 61 PFVGRCCYCTPOSWDKFFNPSPSLGLRNVIIYNETHTRHGLARLSVYLFQIQRDV 120

Qy 121 HKGMFATNVTENVLNSRVSQEAIAEVAALNPDGSAQQSKAVNKKVKKAKRIQLQEMVAT 180

Db 121 HKGMFATNVTENVLNSRVSQEAIAEVAALNPDGSAQQSKAVNKKVKKAKRIQLQEMVAT 180

Qy 181 VSPAMIRLTGWLKLFNSPFFWNIQHKGOLEVMVKAATETNLLPLFLPVHRSHDYLILLT 240

Db 181 VSPAMIRLTGWLKLFNSPFFWNIQHKGOLEVMVKAATETNLLPLFLPVHRSHDYLILLT 240

Qy 241 FILFCHNIKAPYIASGNLNIPIFSTLIHKLGFFIRRLDETDPGRKDVLYRALLHGI 300  
Db 241 FILFCHNIKAPYIASGNLNIPIFSTLIHKLGFFIRRLDETDPGRKDVLYRALLHGI 300  
Qy 301 VELLROQOQFLEIFLEGTRSRGKTSARAGLLSVVVDLTSTNVIPIILLIPVGISYDRII 360  
Db 301 VELLROQOQFLEIFLEGTRSRGKTSARAGLLSVVVDLTSTNVIPIILLIPVGISYDRII 360  
Qy 361 EGHVNGQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQKPVSA 420  
Db 361 EGHVNGQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQKPVSA 420  
Qy 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSC 480  
Db 421 LLSLEQALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSC 480  
Qy 481 AIMSTHIVACLLLYRHRQIDLSLTVEDFFVMKEVLARDFDLGFGNSDEVDVWHAIQLL 540  
Db 481 AIMSTHIVACLLLYRHRQIDLSLTVEDFFVMKEVLARDFDLGFGNSDEVDVWHAIQLL 540  
Qy 541 GNCVTIHTSRNDEFFITPSTTPVSPVFEALNFYSNGVLHVIMEAIACSLYAVLNKRGILG 600  
Db 541 GNCVTIHTSRNDEFFITPSTTPVSPVFEALNFYSNGVLHVIMEAIACSLYAVLNKRGILG 600  
Qy 601 GPTSTPNLISQEQLVKAAASLCVLLSNEGTSISLPCQTFYQVCHETVGKTIQYGLTVAE 660  
Db 601 GPTSTPNLISQEQLVKAAASLCVLLSNEGTSISLPCQTFYQVCHETVGKTIQYGLTVAE 660  
Qy 661 HDQOEDISPLAQOWDKLPEPLSMWSRDEBDESDFGGEQRDCYLKVSQKSHQOPIIF 720  
Db 661 HDQOEDISPLAQOWDKLPEPLSMWSRDEBDESDFGGEQRDCYLKVSQKSHQOPIIF 720  
Qy 721 LQRLGLPALLEAYSSAIFVHNFSGPVPEPEYLOKHLKYLITRERNVAVVAESATYCLVK 780  
Db 721 LQRLGLPALLEAYSSAIFVHNFSGPVPEPEYLOKHLKYLITRERNVAVVAESATYCLVK 780  
Qy 781 NAVKMFKDIGVFKETKQKRVSVLELSSTFLPQCNRQKLLLEYILSFVVL 828  
Db 781 NAVKMFKDIGVFKETKQKRVSVLELSSTFLPQCNRQKLLLEYILSFVVL 828

RESULT 6  
ID ABP69808  
XX  
AC ABP69808;  
XX  
DT 20-JAN-2003 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 1855.  
XX  
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;  
cell-proliferative disorder; neurodegenerative disease; bacterial;  
Parkinson's disease; Alzheimer's disease; autoimmune disease;  
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;  
antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
haemostatic; vulnary; fungicide; antibacterial; virucide; protozoacide;  
antiarthritic.  
XX  
OS Homo sapiens.  
XX  
PN WO200270539-A2.  
XX  
PD 12-SEP-2002.  
XX  
XX 05-MAR-2002; 2002WO-US05095.  
XX  
XX 05-MAR-2001; 2001US-0799451.  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;

PI 102 FILFCHNIKAPYIASGNLNIPIFSTLIHKLGFFIRRLDETDPGRKDVLYRALLHGI 300  
PI 102 FILFCHNIKAPYIASGNLNIPIFSTLIHKLGFFIRRLDETDPGRKDVLYRALLHGI 300  
XX  
DR WPI; 2002-759812/82.  
DR N-PSDB; ABZ12025.  
XX  
XX New polynucleotides comprising sequences assembled from expressed  
sequence tags (ESTs), useful for treating cell-proliferative,  
neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or  
platelet or coagulation disorders -  
PS Claim 9; SEQ ID NO 1855; 1012pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated polynucleotide (I) comprising a  
nucleotide sequence selected from any of 948 sequences  
(ABZ1119-ABZ12066) or their mature protein coding portion, active domain  
coding protein or complementary sequences. The polynucleotides are useful  
for identifying expressed genes or for physical mapping of human genome.  
CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular  
weight markers, as a food supplement, for generating antibodies, in  
medical imaging, screening and diagnostic assays and for treating  
cell-proliferative disorders (cancer), neurodegenerative diseases  
(Parkinson's or Alzheimer's disease), autoimmune diseases (multiple  
sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid  
disorders, platelet or coagulation disorders, wound, burns, incision,  
ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,  
parasitic), arthritis, etc.  
CC Note: The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 705 AA;

Query Match 84.2%; Score 3602; DB 23; Length 705;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 124 MFATNTVENLVNLSRRVQEAIAEVAELNPDGSAQQSKAVNKVKAKRILQBMVATVSP 183  
Db 1 MFATNTVENLVNLSRRVQEAIAEVAELNPDGSAQQSKAVNKVKAKRILQBMVATVSP 60  
Qy 184 AMIRLTGWVLLKLFNSFFWNIIQIHKGQLEVMKAAATETNLPFLPVRSHIDYLLTFFIL 243  
Db 61 AMIRLTGWVLLKLFNSFFWNIIQIHKGQLEVMKAAATETNLPFLPVRSHIDYLLTFFIL 120  
Qy 244 FCHNIKAPYIASGNLNIPIFSTLIHKLGFFIRRLDETDPGRKDVLYRALLHGHVEL 303  
Db 121 FCHNIKAPYIASGNLNIPIFSTLIHKLGFFIRRLDETDPGRKDVLYRALLHGHVEL 180  
Qy 304 LQOQOQFLEIFLEGTRSRGKTSARAGLLSVVVDLTSTNVIPIILLIPVGISYDRIIIEGH 363  
Db 181 LQOQOQFLEIFLEGTRSRGKTSARAGLLSVVVDLTSTNVIPIILLIPVGISYDRIIIEGH 240  
Qy 364 YNGEQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQKSPVSAALLS 423  
Db 241 YNGEQLGPKKNESLWSVARGVIRMLRKNYGCVRVDFAPQPSLKEYLESQKSPVSAALLS 300  
Qy 424 LQOALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSCAIM 483  
Db 301 LQOALLPAILPSRPSDADEGRDTSINESRNATDESRLRLIANLAHILFTASKSCAIM 360  
Qy 484 STHIVACLLLYRHRQIDLSLTVEDFFVMKEVLARDFDLGFGNSDEVDVWHAIQLLGNC 543  
Db 361 STHIVACLLLYRHRQIDLSLTVEDFFVMKEVLARDFDLGFGNSDEVDVWHAIQLLGNC 420  
Qy 544 VTITHTSRNDEFFITPSTTPVSPVFEALNFYSNGVLHVIMEAIACSLYAVLNKRGILGGPT 603  
Db 421 VTITHTSRNDEFFITPSTTPVSPVFEALNFYSNGVLHVIMEAIACSLYAVLNKRGILGGPT 480  
Qy 604 STPPNLISQEQLVKAAASLCVLLSNEGTSISLPCQTFYQVCHETVGKTIQYGLTVAEHDD 663  
Db 481 STPPNLISQEQLVKAAASLCVLLSNEGTSISLPCQTFYQVCHETVGKTIQYGLTVAEHDD 540







capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL016176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX	SQ	Sequence	850 AA;	
XX	Query Match	19.5%;	Score 834; DB 22; Length 850;	
XX	Best Local Similarity	27.5%;	Pred. No. 8.7e-76;	
XX	Matches	230; Conservative	182; Mismatches 313; Indels 110; Gaps 22;	
Qy	42	TVFERSATLKKESLMRKEPFVGRCCYCTPQSDWKFFNPSPISGLRNVIYINETHRH	101	
Db	45	SLFPNLSYRTSYLSKINTIRSDLLVLRSSNEPKTONSAHGIGTQNMLEL----	TPH 100	
Qy	102	RCGLARRLSY---VLFQERDVHKG--MFATVNTENVLNSSLRVOAIAEVAE-----	149	
Db	101	SGQKPKKLPDGVLPFYVAQVVRSEKFEYRQVTEIVHNDVLAIKQAEOITLREORY	160	
Qy	150	-----LNPDGSAQOQSK-----AVNKVKKAKRILQEMVATVSPAMIRLT	189	
Db	161	AKNGHRLLTRSASGQEAQEADEKRGISYQAILRKQORALISILKDMGSTLNGLLAFT	220	
Qy	190	GWVLKLFNSPWNQIHKQLEMYKAATETN--LPILFLPVHRSHIDYLLLTFLPCFN	247	
Db	221	SMILYKLLPFLSGVVTNTKQIEMLTATERSPGTFLFVPLRSHLDYIMVTLITWND	280	
Qy	248	IKAPVIAGNNLNIPIFSLTHLKGFFIRRLDETDPGRKDVLYRALLGHIVELLRQO	307	
Db	281	IRSPVIAAGNNLQIPVFGLLRGLGAFPIKRIKIDPV--EGKDVLYRAALHLYLTHALQO	339	
Qy	308	QFLEIFLEGTRSRSGKTSACAGLSLVVVVDLTSTNTVPIDILIPVGISYDRIIEGHYNGE	367	
Db	340	HNVEFFIEGGRTRTKCPMPKGGILSVIVNAFMDSIFDALLVPVSVNYERLVGDNFVRE	399	
Qy	368	QLGKPKNESLWSVARGVIRMLRKNYGCVRVDFAQFSLKEYLESQSO-----	415	
Db	400	QKGEKKIPESGKATSGIKWALKSNYGLMIDFNEPYSIRELVNSYNIAREDNIAKVY	459	
Qy	416	KPVSAALLSLEQALLPAILPSRPSDAADGRDTSINESNATDESRLRLIANLAHILFT	475	
Db	460	KPSARVLOHQ-----STSLYGTDVVCEEHRN-----LIESIRQVQVFD	499	
Qy	476	ASKCAIMSTHIVACLLYHRQID---LSTLVDFDFVWKEEVLARDFDLFGSGNSRDV	532	
Db	500	CAATSVNMSTNALAFLLLTRFNGAEEQILSEALDDL---RNSLSGCKDIFGSGESSQI	555	
Qy	533	VMAHQLLGNVCVTTHTRSRNDEFFTPSTVPVSELPNFYNGVLVHFIMBAIACSUYA	592	
Db	556	VAYACDLGSLGIVTRSDENGRLVIAKNSVESFIELAYYSNMLTPHPALSSILLTTHS	615	
Qy	593	VL---NKRGLGGPTSTPNLISQOLVKAASLCYLLSNEGTSLSLPCQTFYQVCHETVG	648	
Db	616	LLPETENKEAA-----VSRKKLIDTALENCQIYRYFILNKPTQVLENLLYQQLD	666	
Qy	649	KFIQYQ-ILTVAEHDDQEDISPLAEQOWDKKLPEPLSWRSDDED--EDSDFGEQRDCYL	706	
Db	667	DLISGCVLTKEKLDH---LPGNAE---GRRLANVLAECIDEDGYEDVRGEADEPKLL	719	
Qy	707	KVSGKEHQOFTIFLORLLGPLLEAYSAA---IPVHNFSGVPPEPYQLKLHLYLIRT	763	
Db	720	FASETPSQORYIC---EVLAPFANTYTVTAQSLQILHKNS--MLESEFISFVINDLSKV	774	
Qy	764	ERNVAVYAESATYCLVKNVXWFKDIGYFKETKQKRVSVLESLSTFLPQCNRQKL	818	
Db	775	KRGCIYAESISTDSVRNCLLKEKWSVIEVCNQOQMELISLNTLY--EMGRESL	827	

RESULT 10	
ABG23092	
ID	ABG23092 standard; Protein; 807 AA.
XX	
AC	ABG23092;
XX	
DT	18-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #23083.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder.
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US08631.
XX	
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Dmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	N-PSDB; AAS87279.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity
XX	
Claim 20;	SEQ ID No 53451; 103pp; English.
PS	
XX	
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG0010-ABG30377 represent novel human
CC	diagnostic amino acid sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 807 AA;

Query Match	19.0%;	Score 813; DB 22; Length 807;	
Best Local Similarity	29.0%;	Pred. No. 1.2e-73;	
Matches	233; Conservative	125; Mismatches 299; Indels 146; Gaps 20;	
Qy	56	MSRKRPFVGRCCYCTPQSDWKFFNPSPISGLRNVIYINETHRHGWLARRLSYVLF	115
Db	107	LGRTRPFVGRCCYCTPQSDWKFFNPSPISGLRNVIYINETHRHGWLARRLSYVLF	150
Qy	116	QERDVHKGMEFATNTVNTENVLNSSLRVOAIAEVAELNPDGSAQOQKSAVNVKVKAKRILQ	175

Db 151 ----- 150  
Qy 176 EMVATSPAMIRLTGTVLLKLFNSFFWNIQIHKGQLEWVKAAT-----ETN 221  
Db 151 -----RLFSWALLRVLLKLFUNVQLHKGQKQVOKAAQVQGVQVTSWMLSPQEG 199  
Qy 222 LPLFLFVHRSHIDYLLLTILFELCHNIKAPYIASGNNLNIPFSTLIHKLGGFFIRRLD 281  
Db 200 LPLVLLSTHKTLDGILLPMLLSQGLGVLRVAWDSRACSPALRALLKGLFLPPEAS 259  
Qy 282 ETPDGRKDVLYRALLHGHIVELLRQQQFLFIEFLETRSRSG-KTSCARAGLLSVVVDTLUS 340  
Db 260 LSLDSSEGLLARAVQAVIQLQLVSGQPLIFLEPEPPGALGPLRLSALGQAVGVFVQAVQ 319  
Qy 341 TNVIPDILLIPVGISVDRIEIGHYNGEQLKPKKNESLMSVARGVIRMLRKNGV-----C 395  
Db 320 VGIPLDALLVPVAVTVDLPDPCDIDHASAPL---GLWTGALAVLRSLWSRWGCSHRIC 376  
Qy 396 VRYDFAQPSLKEYLESQSKPVSAALLSLQALLPAILP--SRPSDAADGRDTSINESR 453  
Db 377 SRVHLAQPFSLQEIYVS-ARSCWGRQTLEQLLPVILGQCTAVPDTKEQEWTPTITGPL 435  
Qy 454 NATDESURRLRIANLAEHLIFTASKSCAINTSHIVACLLLYHRHQGIDISTLVEDFFVK 513  
Db 436 LALKEE-DQLLVRLSCHVLSASVSGSSAVMSTAIMATLLLFKHQKGVFLSGLLGEFSWLT 494  
Qy 514 EEWLARDFDLFGFSGNEDVVMHAIQLLGNCTVITHTSRNDEFITPTSTVPSVPELNFYS 573  
Db 495 EEILVRGFDVFGSGQLRSLQHSLSLLRAHVALLRI-ROGDLVAVPQPG-PGLTHLAQLS 552  
Qy 574 NGVLHVFIMEAITACSLIYAVLNKRGILGGPTSTPPN-----LISQEOVLVKAASLCY 624  
Db 553 AELLVFLSVAACAVRGLLAGR-----VPPQGPWELQGLILLSQNELRYQIILLMH 605  
Qy 625 LLNNEGTSILPQCTFYOVCHETVKGFTIQYGLTVAEHDDQEDISPLAQQWQDKLPEPL 684  
Db 606 LLPQDLLLLKPCQSSCYCQEVLDRLITCGLL-VABETPGSRPACDTGRLSRK----L 660  
Qy 685 SWR--SDEEEDS-DEGEQRDCYLVKSQSKHQQTITFLQRLGLLPLEAYSAAAFVHN 741  
Db 661 LWKPSGDFDSDSDDFEAD-GRYFLRSQSQSHCPDFFLFCRLLSPLLKAFQAAPFLR- 718  
Qy 742 FSGVPPEPE--YLOKLHKYILITERNVAVYAESATVCLVKNVAKMFKDGVKPKQKR 799  
Db 719 -QQQLPDTLGYTEQLFQFL-QATAEGEIGFECAAPKASPSAVWTFRLDGLVLQTRSPA 776  
Qy 800 VSVLELSSTFLPQCNKQKLEYI 822  
Db 777 GPRHLGLPLLALDNOEKLRTV 799

RESULT 11  
ABBI1171

AC ABB11171 standard; peptide; 156 AA.

XX ABB11171;

XX 11-JAN-2002 (first entry)

XX Human sn-glycerol-3-P acyltransferase homologue, SEQ ID NO:1541.

XX Human; cytokine; cell proliferation; tissue growth; immunomodulator; growth factor;  
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;  
XX inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;  
XX proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
XX chronic inflammatory condition; proliferative retinopathy;  
XX atherosclerosis; coronary heart disease; arterial ischaemia;  
XX bone disorder; osteoporosis; vascular growth disorder;  
XX tissue regeneration; wound healing; infection; immune disorder;  
XX cell culture; drug screening; gene therapy; antiinflammatory;  
XX antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
XX cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;

KW antifungal; vulnery; antiulcer.

XX Homo sapiens.

PN WO200157188-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US03800.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457740/49.

XX N-PSDB; ABA08415.

Human proteins and DNA encoding sequences useful for preventing,

treating or ameliorating a medical condition in a mammalian subject

e.g. arthritis and cancer -

Claim 20; Page 152; 1963pp; English.

Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity;

haematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a novel human polypeptide of the invention.

Sequence 156 AA;

Query Match 17.2%; Score 738; DB 22; Length 156;

Best Local Similarity 100.0%; Pred. No. 3.6e-67;

Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 RNVIYINETHRRHGLARRLSYVLFQIBRDVHKGFATNTENVLNSRRVQEAIAEVA 148

Db 12 RNVIYINETHRRHGLARRLSYVLFQIBRDVHKGFATNTENVLNSRRVQEAIAEVA 71

QY 149 ELNPDGSAQQOQSKAVNKKKRILOEMVATVSPAMIRLTGWLKLFNSFFWNIQIHK 208  
 |||||  
 Db 72 ELNPDGSAQQOQSKAVNKKKRILOEMVATVSPAMIRLTGWLKLFNSFFWNIQIHK 131  
 |||||  
 QY 209 GOLEMVKAATETNLPPLFLPVHRSH 233  
 |||||  
 Db 132 GOLEMVKAATETNLPPLFLPVHRSH 156  
 |||||

RESULT 12  
 AAY72134  
 ID AAY72134 standard; Protein; 827 AA.  
 XX  
 AC AAY72134;  
 XX  
 DT 24-APR-2001 (first entry)  
 XX  
 DE E. coli glycerol-3-phosphate acyltransferase with ER retention sequence.  
 XX  
 KW Triacylglyceride; TAG; glycerol-3-phosphate acyltransferase; GPAT;  
 XX fatty acid; oil; errs; endoplasmic reticulum retention sequence.  
 XX  
 OS Escherichia coli.  
 XX  
 PN WO200078974-A2.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 20-JUN-2000; 2000WO-CA00738.  
 XX  
 PR 21-JUN-1999; 99US-0139788.  
 PR 16-FEB-2000; 2000US-0182905.  
 XX  
 PA (CANA ) NAT RES COUNCIL CANADA.  
 XX  
 PI Jain RK, Mackenzie SL;  
 DR N-PSDB; AAD02333.  
 XX

Transforming organisms with DNA encoding glycerol-3-phosphate  
 acyltransferase to increase their natural oil and fat levels and alter  
 the fatty acid composition of their triacylglycerides -  
 Claim 20; Page 56-60; 60pp; English.  
 CC The patent discloses a method for increasing the triacylglyceride (TAG)  
 CC content of an organism and/or for modifying the fatty acid composition  
 CC of the TAG by expressing in the organism, a DNA encoding a protein with  
 CC glycerol-3-phosphate acyltransferase GPAT activity. This method is used  
 CC to increase the natural oil and fat levels and to alter the fatty acid  
 CC composition of TAGs in plants and yeast. It can be used to manipulate  
 CC oil synthesis in other organisms such as yeast, other fungus and algae  
 CC for producing commodity and specialty oils. Increasing the oil content  
 CC of feed quality grains reduces the need for adding exogenous fats in the  
 CC diets of animals and birds.  
 CC The present sequence is Escherichia coli glycerol-3-phosphate acyl-  
 CC transferase (GPAT) with endoplasmic reticulum retention sequence (errs).  
 CC The p18+errs gene encoding this sequence is used to modify the type of  
 CC fatty acid at the sn-1 position of triacylglycerides (TAG). This enables  
 CC the production of structured TAGs, in which the fatty acids occupying  
 CC each position may be controlled. It has implications for manipulating  
 CC the fat content in humans and other animals.  
 XX Sequence 827 AA;

Query Match 9.1%; Score 391.5; DB 22; Length 827;  
 Best Local Similarity 24.0%; Pred No. 3.1e-30;  
 Matches 147; Conservative 118; Mismatches 248; Indels 99; Gaps 21;  
 QY 76 DKFFNPSTPSGLRNVIIYNETHRRHGWARRLSYLVFIQERDVHKGMFATNTVNTVLN 135  
 |||||  
 |||||

Db 182 DSFVRFS-PSVSLRRMADEHGTDKTIAQKLA-RVARMHFARQRLAAVGP-RLPARQDLFN 238  
 QY 136 SSRVQEATAEVAALNPDGSAQQOQSKAVN--KVKKAKRILOEMVATVSPAMIRLTGWL 193  
 |||||  
 Db 239 KLLASRAIAKAVED-----EARSKKLSHEKAQONATAIMEETIANFSEYEMIRLTDRL 291  
 |||||  
 QY 194 LKLFNSFFWNIQIHKGOLEMVKAATETNLPPLFLPVHRSHIDYLLLTFLFCHNIKAPYI 253  
 |||||  
 Db 292 GFTWNRLYQGINVHNA--ERVRLAHGDHGLVVPVCHRSHMDYLLLSVLYVHOGVLPPPHI 349  
 |||||  
 QY 254 ASGNLNI-----PIFSTLIHLKLGFFIRRRRLDETPDGRKQVLYRALLHGHIVELLRQOQF 309  
 |||||  
 Db 350 AAGINLNFWPAGPIF----RELGAFFIR---TFKGNK--LYSTVFREYLGELFSRGYS 399  
 |||||  
 QY 310 LEIFLEGRSRSGKTSACAGLLSVVDLTSTNVIPDILIIIPVGISYDRIIE-GHYNGEQ 368  
 |||||  
 Db 400 VEYFVEGGRSRTGLLDPKGTLSMTQAMLRGGTRPTILPIYIGYEHVVEVGYAKEL 459  
 |||||  
 QY 369 LGKPKNESLWSVARGVIRMLRKNYGVVRVDFAPQFSLKEYLES-----OSQKPVSA 422  
 |||||  
 Db 460 RGATKEKESLPQMLRGLSKL--RNLGQGVNFGPMPMLMTYLNQHVDPDWRESIDFI---- 513  
 |||||  
 QY 423 SLEQALLPAILLPSRPSDAADGRTSINESRNADESRLRLIANLAHILFTASKSCAI 482  
 |||||  
 Db 514 ---EAVRPWLTP-----TVNNIAADLMVRINNAGAA 542  
 |||||  
 QY 483 MSTHIVACLLLYRHRQIGIDLSTLVED---FFVMKEEVLARDFDLGFSGNSDEVVMAIQ 538  
 |||||  
 Db 543 NAMNLCTALLASQRSUTREQLTEQLNCYLDLMRNVPYSTDSTVP-SASASELIDHALQ 601  
 |||||  
 QY 539 LLGNCVTITHTSRNDEFFITPSTTVPSVFEINFGVSLVHVFIMEAITACSLYAVLNKRG 598  
 |||||  
 Db 602 M--NKFEVEKDTIGDIILPREQAV----LMTYRNIAHMLVLPMLAAIV----- 647  
 |||||  
 QY 599 LGGTSTPPNLIISOELVVRKAASLCYLLSNGTISLPCQTFYQVCHETVGFQVIGILTV 658  
 |||||  
 Db 648 -----TQHRHISRDVLMEHVNVLYPMLKAEFLFLWRDRDELDPVIDALANEMQROGLITL 701  
 |||||  
 QY 659 AEHDDQEDISPS 670  
 |||||  
 Db 702 --QDELHINPA 711  
 |||||

RESULT 13  
 AAY72133  
 ID AAY72133 standard; Protein; 827 AA.  
 XX  
 AC AAY72133;  
 XX  
 DT 24-APR-2001 (first entry)  
 XX  
 DE Escherichia coli glycerol-3-phosphate acyltransferase (GPAT).  
 XX Triacylglyceride; TAG; glycerol-3-phosphate acyltransferase;  
 KW GPAT; fatty acid; oil.  
 XX  
 OS Escherichia coli.  
 XX  
 PN WO200078974-A2.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 20-JUN-2000; 2000WO-CA00738.  
 XX  
 PR 21-JUN-1999; 99US-0139788.  
 PR 16-FEB-2000; 2000US-0182905.  
 XX  
 PA (CANA ) NAT RES COUNCIL CANADA.  
 XX  
 PI Jain RK, Mackenzie SL;  
 DR WPI; 2001-041428/05.  
 DR N-PSDB; AAD02332.



Query Match 8.5%; Score 362; DB 23; Length 77;  
 Best Local Similarity 92.1%; Pred. No. 5.7e-29;  
 Matches 70; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 155 SAQQQSKAVNKKAKRILOEMVATSPAMIRLTGWLLKLFNSFFWNIQIHKQLEMV 214  
 DB 2 SAQQQSKAIQKVRKRILOEMVATSPGMIRLTGWLLKLFNSFFWNIQIHKQLEMV 61

QY 215 KAATETNLPPLPLPVH 230  
 DB 62 KAATETNLPPLPLPVH 77

RESULT 15  
 ABB60776  
 ID ABB60776 standard; Protein; 724 AA.  
 AC ABB60776;  
 XX

DT 26-MAR-2002 (first entry)  
 XX Drosophila melanogaster polypeptide SEQ ID NO 9120.  
 DE Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 KW Drosophila melanogaster.  
 OS WO200171042-A2.  
 PN 27-SEP-2001.  
 PD

XX 23-MAR-2001; 2001WO-US09231.  
 XX 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX (PEKE ) PE CORP NY.  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656960/75.  
 DR N-PSDB; ABL04879.  
 XX

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX Disclosure; SEQ ID NO 9120; Zipp + Sequence Listing; English.  
 PS The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 724 AA;

Query Match 8.2%; Score 349.5; DB 22; Length 724;  
 Best Local Similarity 20.8%; Pred. No. 5.2e-26;  
 Matches 169; Conservative 135; Mismatches 279; Indels 228; Gaps 31;

QY 117 ERDVHKGMFATVNTVNLN-----SSRQVQAEVAE-----LNP----- 152  
 DB 21 EOEQKTPSAEYMRNFKNIAPGNEASMTREFNPQVAYEFKYNLPKQKQHLRSEKL 80

QY 153 ----DGSAAQQSKAVNKKAKRILOEMVATSPAMIRLTGWLL-----KLFNSFFWNI 204  
 DB 81 RSILEHYAKESGTPKQMERQARALIDIGLDRNMAIRWCGIAITAIGKRICDGFY--- 137  
 QY 205 QIHKQLEMVKAATETNL-----PLLFLPVHRSHIDYLLLTFLFCHNINKAPIASGN 257  
 DB 138 -----VNSASVANVRKMDGKPCVLYLSPSHRSYMDFILMSYICYDYDIEIPGAAAGM 188  
 QY 258 NLNTPI--FSTLHLKGGFFIRRLDETDPGRKDVLYRALLHGHIVELLERQQOF--LEIFLE 315  
 DB 189 DFHSFMFGMTMLRTKGAFMRSPS-----NDELYWDIFREYMTALVANTHIGVEFFIE 242  
 QY 316 CTRSRSGKTSCARAGLLSVVVDTLSTNVIPDILIPVGISYDRIIEGH--YNGEQLGKPKK 374  
 DB 243 CTRSRNFKALVPKIGLLSMALLPYFTGVPDVMIVPVSVAYERVLEEQLFVYELLGVKPK 302  
 QY 375 NESLWSVARGVIRMLR-----KNYGCVRVDFAQPFSLKEYLESQSKPVSALLSLEQALLPA 431  
 DB 303 KES-----TKGFFKALKIIDERFGKMFDFGEPISVKEFF----- 337  
 QY 432 ILPSRPSDAADGDRDTSINESRNATDESRLR-----RLIANLAELHILFTASKSCAINTST 485  
 DB 338 -----GHDSAQRMRAGVGGHLQKLNQOEVELVQLANEIYQQORRIVISTP 385  
 QY 486 HIVACLLYRHRQIGIDLTSLVEDPFVMEKEVLARDFDLGFSGNSEDVVMH---ATQLLGN 542  
 DB 386 NLLS--LYVASQLVAQRSVTLDE-----LARG-----VHLKRIFEQLG- 422  
 QY 543 CVTITHSRNDEFFITPTTVPSPFELNFYNGVLHVFIEMAIACSLYAVLNKRLGGP 602  
 DB 423 -----AHVS-----TTPSSIKADVIDAVEIHSNHLF-----ATGRL 454  
 QY 603 TSTP---PNLISQBLVR-KAASLCYLLSNEGTSILPCQTFYQVCHETVGKFIQVGLTV 658  
 DB 455 QFTPMKQLAQLDVKRLKAHALCPQTMVAVPMLALQLYNCPMFLARPAYLLAAL 514  
 QY 659 AEHDDQEDISPLAEQQWKKLPPLSPWSRDEDEDSDFGEQRDCYLKVSQSKHQOFI 718  
 DB 515 KEQKNQKQSTDIS---YDASL---CALHAHVTTMDALFOHE-----PIESNREAAEF 563  
 QY 719 TFLQRLGPG-LLEAYSSAIPVHN-----HKVLIIRTRNVAVYAESATYCLVKNVAKMFKDIG 744  
 DB 564 THLQLLDERVVEVETSGRINVDNECSHVILAALAPFLCLYQLVVTLRKIPLELEFSN 623  
 QY 745 P---VPPEPYLOKL-----P---CALHAHVTTMDALFOHE-----PIESNREAAEF 790  
 DB 624 KELLVRVQHVQEQVLLQOPGASASHVHPYCLALDNLNIAIYALIORGYLVKS-----RDSG 678  
 QY 791 VFKETKQKRVSVLELSSTFLPQCNRQKLEY 821  
 DB 679 QMKIATPGKCLRELETQLLEYCOLMPPAQY 709

Search completed: January 10, 2004, 20:18:31  
 Job time : 62 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2004, 04:06:03 ; Search time 574.937 Seconds  
(without alignments)  
15003.329 Million cell updates/sec

Title: US-09-935-290-3  
Perfect score: 2484  
Sequence: 1 atggatgaatgcactgac.....ttctgagtttctgtgtgtg 2484

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2276164 seqs, 1736306516 residues  
Total number of hits satisfying chosen parameters: 4552328

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2484	100.0	2484	11	US-09-935-290-3
2	2484	100.0	3003	11	US-09-935-290-1
3	2484	100.0	4033	13	US-09-814-353-20793
4	2480.8	99.9	3210	13	US-10-288-252-26
5	1879.2	75.7	2646	10	US-09-917-800A-1648
6	581.6	23.4	612	10	US-09-833-381-1961
7	437	17.6	482	11	US-09-918-995-21644
8	385.4	15.5	399	11	US-09-918-995-7822
9	348.4	14.0	440	10	US-09-833-381-1964
10	346.6	14.0	396	10	US-09-833-381-1959
11	342.2	13.8	429	10	US-09-960-352-7384
12	337.6	13.6	407	10	US-09-960-352-8821
13	331.8	13.4	868	10	US-09-833-381-1958
14	302	12.2	418	10	US-09-960-352-742
15	268.6	10.8	379	13	US-09-814-353-16182

c	16	233	9.4	303	10	US-09-960-352-13086	Sequence 13086, A
c	17	209.6	8.4	264	10	US-09-960-352-6239	Sequence 6239, Ap
c	18	207.2	8.3	244	13	US-09-814-353-3482	Sequence 3482, Ap
c	19	207.2	8.3	244	13	US-09-814-353-9798	Sequence 9798, Ap
c	20	172.6	6.9	387	10	US-09-960-352-7634	Sequence 7634, Ap
c	21	129	5.2	308	10	US-09-833-381-1963	Sequence 1963, Ap
c	22	109	4.4	2682	13	US-10-288-252-27	Sequence 27, Appl
c	23	109	4.4	2755	13	US-10-288-690-31	Sequence 4061, Ap
c	24	99.4	4.0	501	10	US-09-960-352-4061	Sequence 297935, A
c	25	75.2	3.0	467	13	US-10-027-632-297935	Sequence 297935, A
c	26	75.2	3.0	467	13	US-10-027-632-297936	Sequence 297936, A
c	27	75.2	3.0	467	14	US-10-027-632-297935	Sequence 297935, A
c	28	75.2	3.0	467	14	US-10-027-632-297936	Sequence 297936, A
c	29	75.2	3.0	634	13	US-10-027-632-41737	Sequence 41737, A
c	30	75.2	3.0	634	13	US-10-027-632-41738	Sequence 41738, A
c	31	75.2	3.0	634	14	US-10-027-632-41737	Sequence 41737, A
c	32	75.2	3.0	634	14	US-10-027-632-41738	Sequence 41738, A
c	33	62	2.5	442	10	US-09-833-381-1965	Sequence 1965, Ap
c	34	59.2	2.4	469	13	US-10-027-632-62086	Sequence 62086, A
c	35	59.2	2.4	469	13	US-10-027-632-62087	Sequence 62087, A
c	36	59.2	2.4	469	13	US-10-027-632-63434	Sequence 63434, A
c	37	59.2	2.4	469	13	US-10-027-632-63435	Sequence 63435, A
c	38	59.2	2.4	469	14	US-10-027-632-62086	Sequence 62086, A
c	39	59.2	2.4	469	14	US-10-027-632-62087	Sequence 62087, A
c	40	59.2	2.4	469	14	US-10-027-632-63434	Sequence 63434, A
c	41	59.2	2.4	469	14	US-10-027-632-63435	Sequence 63435, A
c	42	57.6	2.3	2262	12	US-10-369-493-33072	Sequence 33072, A
c	43	56.4	2.3	450	10	US-09-833-381-1960	Sequence 1960, Ap
c	44	56	2.3	2262	12	US-10-369-493-32927	Sequence 32927, A
c	45	56	2.3	2613	12	US-10-369-493-41238	Sequence 41238, A

ALIGNMENTS

RESULT 1  
US-09-935-290-3  
; Sequence 3, Application US/09935290  
; Publication No. US20030044948A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller-Libermann, Rosana  
; TITLE OF INVENTION: 56919, A NOVEL HUMAN ACYLTRANSFERASE AND USES THEREOF  
; FILE REFERENCE: MNI-186  
; CURRENT APPLICATION NUMBER: US/09/935,290  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/226,509  
; PRIOR FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2484  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-935-290-3

Query Match	100.0%	Score 2484	DB 11	Length 2484
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2484	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	ATGGATGAATCTGCACCTGACCCCTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCA	60	
Db	1	ATGGATGAATCTGCACCTGACCCCTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCA	60	
Qy	61	GAATACAGTGTGTGTCATGTAAACACACAAAGTGAAGTGGGTGAGTGGCTTTAGA	120	
Db	61	GAATACAGTGTGTGTCATGTAAACACACAAAGTGAAGTGGGTGAGTGGCTTTAGA	120	
Qy	121	CCACCGCTTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTTAATGAGTCGGAAGG	180	
Db	121	CCACCGCTTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTTAATGAGTCGGAAGG	180	
Qy	181	CCATTTGTGAAGATGTTGTTACTCTCGACTCCCGAGCTGGCAAAATTTTCAAC	240	
Db	181	CCATTTGTGAAGATGTTGTTACTCTCGACTCCCGAGCTGGCAAAATTTTCAAC	240	



Db 181 CCATTTGTTGGAAGATGTTGTTACTCTCTGCACTCCCGAGAGCTGGGCAAAATTTTCAAC 240  
Qy 241 CCAGTATCCGCTCTTTGGGTTTTCGGAATGTTATTTATATCAATGAAATCTCACACAAG 300  
Db 241 CCAGTATCCGCTCTTTGGGTTTTCGGAATGTTATTTATATCAATGAAATCTCACACAAG 300  
Qy 301 CACCGGATGGCTTTCGGAAGACGCTTTCTTACGCTCTTTTATTTCAAGACGAGATGTG 360  
Db 301 CACCGGATGGCTTTCGGAAGACGCTTTCTTACGCTCTTTTATTTCAAGACGAGATGTG 360  
Qy 361 CATAGGGCATGTTTGGCCACAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 420  
Db 361 CATAGGGCATGTTTGGCCACAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 420  
Qy 421 GAGCAATTCGAGAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCCGACGCAATCA 480  
Db 421 GAGCAATTCGAGAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCCGACGCAATCA 480  
Qy 481 AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTAAAGGATTTCTTCAAGAAATGGTTGCCACT 540  
Db 481 AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTAAAGGATTTCTTCAAGAAATGGTTGCCACT 540  
Qy 541 GTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGCTAAACCTGTTCAACAGCTTC 600  
Db 541 GTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGCTAAACCTGTTCAACAGCTTC 600  
Qy 601 TTTTGGAACTTCAAAATTCACAAAGTCAACTGAGATGGTTAAAGCTGCAACTGAGAG 660  
Db 601 TTTTGGAACTTCAAAATTCACAAAGTCAACTGAGATGGTTAAAGCTGCAACTGAGAG 660  
Qy 661 AATTGGCGCTCTGTTTCTACCAAGTTTACAGATGCCATATTTGACTATCTGCTGCTCACT 720  
Db 661 AATTGGCGCTCTGTTTCTACCAAGTTTACAGATGCCATATTTGACTATCTGCTGCTCACT 720  
Qy 721 TTCAATCTCTCTGCTCCATAACATCAAGACCAATACATGCTTCAAGCAATATCTCAAC 780  
Db 721 TTCAATCTCTCTGCTCCATAACATCAAGACCAATACATGCTTCAAGCAATATCTCAAC 780  
Qy 781 ATCCCAATCTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTCAACGACGAGGCTC 840  
Db 781 ATCCCAATCTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTCAACGACGAGGCTC 840  
Qy 841 GATGAACACACAGATGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db 841 GATGAACACACAGATGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Qy 901 GTTGAATTAATCTTGACAGCAGCAATCTTGAGATCTTCTGGAAGTCAACGCTTCTAG 960  
Db 901 GTTGAATTAATCTTGACAGCAGCAATCTTGAGATCTTCTGGAAGTCAACGCTTCTAG 960  
Qy 961 AGTGGAAAACCTCTTGCTCGGCGAGGACTTTTGTTCAGTTGGGTAGATCTCTGCT 1020  
Db 961 AGTGGAAAACCTCTTGCTCGGCGAGGACTTTTGTTCAGTTGGGTAGATCTCTGCT 1020  
Qy 1021 ACCAATGCTATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATCGCAATATC 1080  
Db 1021 ACCAATGCTATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATCGCAATATC 1080  
Qy 1081 GAAGGTCACTACAAATGTGTGAACAACTGGGCAACCTAAGAAAGATGAGAGCTGTGGAGT 1140  
Db 1081 GAAGGTCACTACAAATGTGTGAACAACTGGGCAACCTAAGAAAGATGAGAGCTGTGGAGT 1140  
Qy 1141 GTAGCAGAGGTGTTATAGATGTTACGAAATACTATGGTTGTGTCGAGTGGATTTT 1200  
Db 1141 GTAGCAGAGGTGTTATAGATGTTACGAAATACTATGGTTGTGTCGAGTGGATTTT 1200  
Qy 1201 GCACAGCATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAAGAAACCGGTGCTGCT 1260  
Db 1201 GCACAGCATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAAGAAACCGGTGCTGCT 1260  
Qy 1261 CTACTTTTCCCTGGAGCAAGCTTGTATACCACTATCTCTTCAAGACCCAGTGATGCT 1320  
Db 1261 CTACTTTTCCCTGGAGCAAGCTTGTATACCACTATCTCTTCAAGACCCAGTGATGCT 1320

Qy 1321 GCTGATGAAGGTAGAGACACGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db 1321 GCTGATGAAGGTAGAGACACGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Qy 1381 CGAAGAGGTTGATTTGCAAAATCTGGCTGAGCATATTTCTATTCTGCTAGCAAGTCTGT 1440  
Db 1381 CGAAGAGGTTGATTTGCAAAATCTGGCTGAGCATATTTCTATTCTGCTAGCAAGTCTGT 1440  
Qy 1441 GCATTTATCTCCACACACATTTGCTGCTGCTGCTCTCTACAGACACAGGAGGAAAT 1500  
Db 1441 GCATTTATCTCCACACACATTTGCTGCTGCTGCTCTCTACAGACACAGGAGGAAAT 1500  
Qy 1501 GATCTCTCCACATTTGGTTCGAAGACTTCTTTGTGATGAAAGAGAACTCTGCTGCTGAT 1560  
Db 1501 GATCTCTCCACATTTGGTTCGAAGACTTCTTTGTGATGAAAGAGAACTCTGCTGCTGAT 1560  
Qy 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGCTGCTG 1620  
Db 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAAGATGTAGTAATGCATGCCATACAGCTGCTG 1620  
Qy 1621 GGAAATTTGTTGATCACAATCACCCACACTAGCAGGAACGATGATTTTATCACCCACG 1680  
Db 1621 GGAAATTTGTTGATCACAATCACCCACACTAGCAGGAACGATGATTTTATCACCCACG 1680  
Qy 1681 ACAACTGTCCCCTCAGTCTTTCGAACTTCTACAGCAATGGGGTACTTCAATGCTTT 1740  
Db 1681 ACACTGTCCCCTCAGTCTTTCGAACTTCTACAGCAATGGGGTACTTCAATGCTTT 1740  
Qy 1741 ATCATGGAGGCCATCATAGCTTTGAGCCCTTTATGCAATGCTGAAAGAGGAGTCTGGG 1800  
Db 1741 ATCATGGAGGCCATCATAGCTTTGAGCCCTTTATGCAATGCTGAAAGAGGAGTCTGGG 1800  
Qy 1801 GGTCCCACTAGCACCCCACTAACCCTGATCAGCAGGAGCAGCTGGTGGAGGCGGCC 1860  
Db 1801 GGTCCCACTAGCACCCCACTAACCCTGATCAGCAGGAGCAGCTGGTGGAGGCGGCC 1860  
Qy 1861 AGCTGTGCTCTCTCTCCAATGAAGCACCCTCTCACTGCTTGGCCAGACATTTTAC 1920  
Db 1861 AGCTGTGCTCTCTCTCCAATGAAGCACCCTCTCACTGCTTGGCCAGACATTTTAC 1920  
Qy 1921 CAACTGCTGCCATGAAACAGTAGAGAAAGTTTATCCAGATGCGCATTTTACAGTGGCAGAG 1980  
Db 1921 CAACTGCTGCCATGAAACAGTAGAGAAAGTTTATCCAGATGCGCATTTTACAGTGGCAGAG 1980  
Qy 1981 CAGATGACCCAGGAAGATATCAGTCTAGTCTTGTGAGCAGCAGTGGGACAGAGCTT 2040  
Db 1981 CAGATGACCCAGGAAGATATCAGTCTAGTCTTGTGAGCAGCAGTGGGACAGAGCTT 2040  
Qy 2041 CCAGAACCTTTGTTTGGAGAAAGTGAAGAGATGAAGACAGTCACTTTTGGGAGGAA 2100  
Db 2041 CCAGAACCTTTGTTTGGAGAAAGTGAAGAGATGAAGACAGTCACTTTTGGGAGGAA 2100  
Qy 2101 CAGCGAGATTGCTTACCTGAAGGTGAGCCAAATCCAAGGAGCACCAGCAGTTTATCACCTTC 2160  
Db 2101 CAGCGAGATTGCTTACCTGAAGGTGAGCCAAATCCAAGGAGCACCAGCAGTTTATCACCTTC 2160  
Qy 2161 TTACAGACATCTCTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTAC 2220  
Db 2161 TTACAGACATCTCTTGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTAC 2220  
Qy 2221 AACTTCAGTGGTCTCTGTTCCAGAACTGAGTATCTGCAAAAGTTGCAAAATACCTAAATA 2280  
Db 2221 AACTTCAGTGGTCTCTGTTCCAGAACTGAGTATCTGCAAAAGTTGCAAAATACCTAAATA 2280  
Qy 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTSCCACAATTTGCTTGTGAAG 2340  
Db 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTSCCACAATTTGCTTGTGAAG 2340  
Qy 2341 AATGCTGTGAAATGTTTAAAGATATTTGGGTTTTCAGGAGACCCAAACAAAGAGAGTG 2400  
Db 2341 AATGCTGTGAAATGTTTAAAGATATTTGGGTTTTCAGGAGACCCAAACAAAGAGAGTG 2400

QY 2401 TCTGTTTGTAGAACTGAGCAGCACTTTTCTACCTCAATCAATGCAACCGCAGAAAAAATCTCTAGAA 2460  
DB 2401 TCTGTTTGTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGCAGAAAAAATCTCTAGAA 2460  
QY 2461 TATATTTCTGAGTTTGTGCTGCTG 2484  
DB 2461 TATATTTCTGAGTTTGTGCTGCTG 2484

RESULT 2

US-09-935-290-1  
; Sequence 1, Application US/09935290  
; Publication No. US20030044948A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller-Libermann, Rosana  
; TITLE OF INVENTION: 56919, A NOVEL HUMAN ACYLTRANSFERASE AND USES THEREOF  
; FILE REFERENCE: MNI-186  
; CURRENT APPLICATION NUMBER: US/09/935,290  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/226,509  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3003  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (341)...(2827)  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(3003)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-935-290-1

Query Match 100.0%; Score 2484; DB 11; Length 3003;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGATGAATCTGCACTGACCTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCA 60  
DB 341 ATGGATGAATCTGCACTGACCTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCA 400  
QY 61 GAATACAGTGTGGTTCGATGTAAGCACACACAGTGAAGAAATGGGTGAGTGGCTTTAGA 120  
DB 401 GAATACAGTGTGGTTCGATGTAAGCACACAGTGAAGAAATGGGTGAGTGGCTTTAGA 460  
QY 121 CCCACCGTCTCAGATCTGCACTTTAAATGGAAGAAAGCCTTAATGAGTCGGAAGG 180  
DB 461 CCCACCGTCTCAGATCTGCACTTTAAATGGAAGAAAGCCTTAATGAGTCGGAAGG 520  
QY 181 CCATTGTTGGAAGATGTTGTTACTCTGCTGCTCCAGAGTGGGACAAATTTTCAAC 240  
DB 521 CCATTGTTGGAAGATGTTGTTACTCTGCTGCTCCAGAGTGGGACAAATTTTTCAC 580  
QY 241 CCCAGTATCCCGTCTTTGGGTTTGGGAAATGTTATTTATATCAATGAAACTCACACAAG 300  
DB 581 CCCAGTATCCCGTCTTTGGGTTTGGGAAATGTTATTTATATCAATGAAACTCACACAAG 640  
QY 301 CACCGGAGTGGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGAGCAGATGTG 360  
DB 641 CACCGGAGTGGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGAGCAGATGTG 700  
QY 361 CATAGGCGATGTTTGGCACCAGTGTGATGAAATGCTGAACAGCAGTAGAGTACAA 420  
DB 701 CATAGGCGATGTTTGGCACCAGTGTGATGAAATGCTGAACAGCAGTAGAGTACAA 760  
QY 421 GAGGCAATTGCAAGTGGCTCTCAATTTAAACCTGTGTTCTGCCCAGCAGCAATCA 480  
DB 761 GAGGCAATTGCAAGTGGCTCTCAATTTAAACCTGTGTTCTGCCCAGCAGCAATCA 820  
QY 481 AAAGCCGTTAAACAGTGAAGAAAGCTAAAGGATTTCTTCAAGAAATGTTGCCACT 540  
DB 1901 TTTGACCTGGGTTCTCAGGAAATTCAGAGATGTAGTAATGATGCTATGCCATACAGCTGCTG 1960

DB 821 AAAGCCGTTAAACAGTGAAGAAAGCTAAAGGATTTCTTCAAGAAATGGTTGCCACT 880  
QY 541 GTCTCACCGCAATGATCAGACTGACCTGGGTGGTGTCTGCTTAAACTGTTTCAACAGCTTC 600  
DB 881 GTCTCACCGCAATGATCAGACTGACCTGGGTGGTGTCTGCTTAAACTGTTTCAACAGCTTC 940  
QY 601 TTTTGGAAACATTTCAAAATTCACAAAGGTCAACTTGCAGATGTTTAAAGCTGCAACTGAGACG 660  
DB 941 TTTTGGAAACATTTCAAAATTCACAAAGGTCAACTTGCAGATGTTTAAAGCTGCAACTGAGACG 1000  
QY 661 AATTTGGCGCTTCTGTTTCTTACAGATTCATAGATCCCATATTTGACTATCTGCTCACT 720  
DB 1001 AATTTGGCGCTTCTGTTTCTTACAGATTCATAGATCCCATATTTGACTATCTGCTCACT 1060  
QY 721 TTCAATCTCTTCTGCCATAACATCAAGACCATACATTTGCTTCCAGGCAATAATCTCAAC 780  
DB 1061 TTCAATCTCTTCTGCCATAACATCAAGACCATACATTTGCTTCCAGGCAATAATCTCAAC 1120  
QY 781 ATCCCAATCTTCACTACCTTGATCCATAAGCTTGGGGCTTCTTATAGAGCTTTGCTCCATGGGCATATA 900  
DB 1121 ATCCCAATCTTCACTACCTTGATCCATAAGCTTGGGGCTTCTTATAGAGCTTTGCTCCATGGGCATATA 1240  
QY 841 GATCAAAACACAGATGACGAGCAATCTTGTGCTCGGCGAGGACTTTTGTAGAGCTTTGCTCCATGGGCATATA 900  
DB 1181 GATCAAAACACAGATGACGAGCAATCTTGTGCTCGGCGAGGACTTTTGTAGAGCTTTGCTCCATGGGCATATA 1240  
QY 901 GTTCAATTTACTTTCAGACAGCAATCTTGTGCTCGGCGAGGACTTTTGTAGAGCTTTGCTCCATGGGCATATA 900  
DB 1241 GTTCAATTTACTTTCAGACAGCAATCTTGTGCTCGGCGAGGACTTTTGTAGAGCTTTGCTCCATGGGCATATA 900  
QY 961 AGTGGAAAAACCTCTTGTGCTCGGCGAGGACTTTTGTAGAGCTTTGCTCCATGGGCATATA 900  
DB 1301 AGTGGAAAAACCTCTTGTGCTCGGCGAGGACTTTTGTAGAGCTTTGCTCCATGGGCATATA 900  
QY 1021 ACCAATGTCTCCAGACATCTTCAATATACCTGTTGGAATCTCTATGATCGCATATC 1080  
DB 1361 ACCAATGTCTCCAGACATCTTCAATATACCTGTTGGAATCTCTATGATCGCATATC 1420  
QY 1081 GAAGGTCACTACAATGTTGAACAACTGGGCAAACTTAAGAAAGTGAAGAGCTGTGGAGT 1140  
DB 1421 GAAGGTCACTACAATGTTGAACAACTGGGCAAACTTAAGAAAGTGAAGAGCTGTGGAGT 1480  
QY 1141 GTAGCAAGAGGTGTTATTAGAAATTTACGAAAAAATATGTTGTTGTTCCGAGTGGATTTT 1200  
DB 1481 GTAGCAAGAGGTGTTATTAGAAATTTACGAAAAAATATGTTGTTGTTCCGAGTGGATTTT 1540  
QY 1201 GCACAGCCATTTTCTTAAAGGAAATTTTGAAGAGCAAAAGTCAAGAACCCGGTGTCTGCT 1260  
DB 1541 GCACAGCCATTTTCTTAAAGGAAATTTTGAAGAGCAAAAGTCAAGAACCCGGTGTCTGCT 1600  
QY 1261 CTACTTTCCCTGGAGCAAGCGTTGTTACCGAGTATCTCTTCAAGACCCAGTGTGCT 1320  
DB 1601 CTACTTTCCCTGGAGCAAGCGTTGTTACCGAGTATCTCTTCAAGACCCAGTGTGCT 1660  
QY 1321 GCTGATGAAGGTAGAGACAGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
DB 1661 GCTGATGAAGGTAGAGACAGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1720  
QY 1381 CGAAGAGGTGTTGCAAAATCTGGCTGAGCATATTTCTTACTGTCTGCTGCTGCTGCTGCT 1440  
DB 1721 CGAAGAGGTGTTGCAAAATCTGGCTGAGCATATTTCTTACTGTCTGCTGCTGCTGCTGCT 1780  
QY 1441 GCCATTAATGTCACACATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
DB 1781 GCCATTAATGTCACACATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1840  
QY 1501 GATCTCTCCATTTGTCGAAAGACTTTCTTGTGATGAAAGAGGAAAGTCTCGCTCGTAT 1560  
DB 1841 GATCTCTCCATTTGTCGAAAGACTTTCTTGTGATGAAAGAGGAAAGTCTCGCTCGTAT 1900  
QY 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAGATGTAGTAATGATGCTATGCCATACAGCTGCTG 1620  
DB 1901 TTTGACCTGGGTTCTCAGGAAATTCAGAGATGTAGTAATGATGCTATGCCATACAGCTGCTG 1960

1621 GGAATTTGTGTCACATCACCACCTAGCAGGAACGATGAGTTTATTCACCCCGAGC 1680  
1661 GGAATTTGTGTCACATCACCACCTAGCAGGAACGATGAGTTTATTCACCCCGAGC 2020  
1681 ACAATGTCCTCATCAGTCTTTCGAACCTCACTTCTACAGCAATGGGGTACTTCTTT 1740  
1701 ACAATGTCCTCATCAGTCTTTCGAACCTCACTTCTACAGCAATGGGGTACTTCTTT 2080  
1741 ATCATGAGGCGCATATAGTCTGAGCCTTTATGCACTTCTGCAATGAGGGAAGTGGG 1800  
1761 ATCATGAGGCGCATATAGTCTGAGCCTTTATGCACTTCTGCAATGAGGGAAGTGGG 2140  
1801 GGTCCCACTAGCAGCCACCTCACTGATCAGCAGGAGCAGTGGTGGGAAGGCGGCC 1860  
1821 GGTCCCACTAGCAGCCACCTCACTGATCAGCAGGAGCAGTGGTGGGAAGGCGGCC 2200  
1861 AGCTGTGCTACCTTCTTCCAAATGAAGGACCATCTCACTGCTTGCAGACATTTTAC 1920  
1901 AGCTGTGCTACCTTCTTCCAAATGAAGGACCATCTCACTGCTTGCAGACATTTTAC 2260  
1921 CAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTTACAGTGGCAG 1980  
1941 CAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTTACAGTGGCAG 2320  
1981 CACGATCAGCAGGAAGATATCAGTCTTCTGCTGAGCAGCAGTGGGACAAAGAGCTT 2040  
2001 CACGATCAGCAGGAAGATATCAGTCTTCTGCTGAGCAGCAGTGGGACAAAGAGCTT 2380  
2041 CCAGAACCTTTGCTTGGAGAGTGTATGAAAGAGATGAAGACAGTGTGCTTGGGGAGAA 2100  
2061 CCAGAACCTTTGCTTGGAGAGTGTATGAAAGAGATGAAGACAGTGTGCTTGGGGAGAA 2440  
2101 CAGGAGATGCTACCTGAAAGTGTAGCCAAATCCAAAGGAGCAGCAGTGTATCACCTTC 2160  
2121 CAGGAGATGCTACCTGAAAGTGTAGCCAAATCCAAAGGAGCAGCAGTGTATCACCTTC 2500  
2161 TTACAGAGATCCTTGGGCTTGTCTGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2220  
2181 TTACAGAGATCCTTGGGCTTGTCTGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2560  
2221 AACTTTCAGTGTCTGTTTCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTAATA 2280  
2241 AACTTTCAGTGTCTGTTTCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTAATA 2620  
2281 ACCAGAACAGAAAGATTTGCAAGTATATGCTGAGAGTGCACATATGCTTGTGAAG 2340  
2301 ACCAGAACAGAAAGATTTGCAAGTATATGCTGAGAGTGCACATATGCTTGTGAAG 2680  
2341 AATGCTGTGAAATGTTTAAGGATATTTGAGGATTTTCAAGGAGACCAAAAGAGAGTG 2400  
2361 AATGCTGTGAAATGTTTAAGGATATTTGAGGATTTTCAAGGAGACCAAAAGAGAGTG 2740  
2401 TCTGTTTATAGAACGTGAGCAGCCTTTTCTACCTCAATGCAACCGACCAAAACTTTAGAA 2460  
2421 TCTGTTTATAGAACGTGAGCAGCCTTTTCTACCTCAATGCAACCGACCAAAACTTTAGAA 2800  
2461 TATATTCAGTGTGTTGTTGCTG 2484  
2501 TATATTCAGTGTGTTGTTGCTG 2824

RESULT 3  
US-09-814-353-20793  
; Sequence 20793, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Pamela  
; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER

FILE REFERENCE: MRI-006B  
CURRENT APPLICATION NUMBER: US/09/814,353  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/191,031  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: US 60/207,124  
PRIOR FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: US 60/211,940  
PRIOR FILING DATE: 2000-06-15  
PRIOR APPLICATION NUMBER: US 60/216,820  
PRIOR FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: US 60/220,661  
PRIOR FILING DATE: 2000-07-25  
PRIOR APPLICATION NUMBER: US 60/257,672  
NUMBER OF SEQ ID NOS: 22037  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 20793  
LENGTH: 4033  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1, 2, 4030, 4031, 4032, 4033  
OTHER INFORMATION: n = A,T,C or G  
US-09-814-353-20793  
Query Match 100.0%; Score 2484; DB 13; Length 4033;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGATGAATCTGCACTGACCCCTGGTACAATAGATGTTTCTTATCTGCCACATTCATCA 60  
DB 341 ATGGATGAATCTGCACTGACCCCTGGTACAATAGATGTTTCTTATCTGCCACATTCATCA 400  
QY 61 GAATCAGTGTGTCGATGTAAGCACACAACTGAGGAATGGGGTGTGCTGCTTTTGA 120  
DB 401 GAATCAGTGTGTCGATGTAAGCACACAACTGAGGAATGGGGTGTGCTGCTTTTGA 460  
QY 121 CCACCGCTTCAGATCTGCACTTTAAATGGAAGAAAGCCTAATGATCGGAAAGG 180  
DB 461 CCACCGCTTCAGATCTGCACTTTAAATGGAAGAAAGCCTAATGATCGGAAAGG 520  
QY 181 CCATTTGTTGGAAGATGTTTGTCTCTGCACTCCCGAGCTGGGACAAATTTTCAAC 240  
DB 521 CCATTTGTTGGAAGATGTTTGTCTCTGCACTCCCGAGCTGGGACAAATTTTCAAC 580  
QY 241 CCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATCAATGAAACTCACACAAG 300  
DB 581 CCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATCAATGAAACTCACACAAG 640  
QY 301 CACCGGATGCTTGGCAAGAGCCCTTCTTACGTTCTTTTATTCAGAGCGAGATGTG 360  
DB 641 CACCGGATGCTTGGCAAGAGCCCTTCTTACGTTCTTTTATTCAGAGCGAGATGTG 700  
QY 361 CATAGGGCATGTTTGGCCACCAATGCTGCTGAAATGCTGCTGAAAGCTGAGTACAA 420  
DB 701 CATAGGGCATGTTTGGCCACCAATGCTGCTGAAATGCTGCTGAAAGCTGAGTACAA 760  
QY 421 GAGGCAATTCAGAGAGTGGCTGCTGAATTTAAACCCCTGATGTTTCTGCCAGCAATCA 480  
DB 761 GAGGCAATTCAGAGAGTGGCTGCTGAATTTAAACCCCTGATGTTTCTGCCAGCAATCA 820  
QY 481 AAGCGGTTAAACAAAGTGAAGAAAGCTTAAAGGATTTCTCAAGAAATGTTGCGACT 540  
DB 821 AAGCGGTTAAACAAAGTGAAGAAAGCTTAAAGGATTTCTCAAGAAATGTTGCGACT 880  
QY 541 GTCTCACCGCAATGATCAGACTGACTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
DB 881 GTCTCACCGCAATGATCAGACTGACTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCT 940  
QY 601 TTTTGAACATTCAAATTCACAAAGCTCAACTTGAGATGTTTAAAGCTGCAACTGAGAG 660

941	T T T T G G A A C A T T C A A A T T C A C A A A G G T C A A C T T G A G A T G G T T A A G C T G C A A C T G A G A C G	1000
661	A A T T T G C C G G T C T G T T T C T A C C A G T T C A T A G A T C C C A T A T T G A C T A T C T G C T G C T C A C T	720
1001	A A T T T G C C G G T C T G T T T C T A C C A G T T C A T A G A T C C C A T A T T G A C T A T C T G C T G C T C A C T	1060
721	T T C A T T C T C T T C G C C A T A C A T C A A A G C A C C A T A C A T T G T T C A G G C A A T A A T C T C A A C	780
1061	T T C A T T C T C T T C T G C C A T A C A T C A A A G C A C C A T A C A T T G C T T C A G G C A A T A A T C T C A A C	1120
781	A T C C C A A T C T T C A G T A C C T T G A T C C A T A A G C T T G G G G C T T C T T C A T A C A C A G A A G G C T C	840
1121	A T C C C A A T C T T C A G T A C C T T G A T C C A T A A G C T T G G G G C T T C T T C A T A C A C A G A A G G C T C	1180
841	G A T G A A A C A C C A G A T G G A C G A A A G A T G T T C T C T A T A G A G T T T G C T C C A T G G C C A T A T A	900
1181	G A T G A A A C A C C A G A T G G A C G A A A G A T G T T C T C T A T A G A G T T T G C T C C A T G G C C A T A T A	1240
901	G T T G A A T T A C T T C G A C A G C A G C A A T T C T T G G A G A T C T T C C T G G A A G G C A C A C G T T C T A G G	960
1241	G T T G A A T T A C T T C G A C A G C A G C A A T T C T T G G A G A T C T T C C T G G A A G G C A C A C G T T C T A G G	1300
961	A G T G G A A A A A C C T C T T G T G C T C G G G C A G A G A C T T T T G N C A G T G T G G T A G A T A C T C T G T C T	1020
1301	A G T G G A A A A A C C T C T T G T G C T C G G C A G A G A C T T T T G T C A G T G T G G T A G A T A C T C T G T C T	1360
1021	A C C A A T G T C A T C C C A G A C A C T T T G A T A A T A C C T G T T G G A A T C T C C T A T G A T C G C A T T A T C	1080
1361	A C C A A T G T C A T C C C A G A C A C T T T G A T A A T A C C T G T T G G A A T C T C C T A T G A T C G C A T T A T C	1420
1081	G A A G G T C A C T A C A A T G T G T A A C A A C T C G G C A A A C C T A A G A A G A A T A G A G C C T G T G G A G T	1140
1421	G A A G G T C A C T A C A A T G T G T A A C A A C T C G G C A A A C C T A A G A A G A A T A G A G C C T G T G G A G T	1480
1141	G T A G C A G A G G T G T T A T T A G A A T G T T A C G A A A A A A C T A T G G T T G T G T C C A G T G E A T T T T	1200
1481	G T A G C A G A G G T G T T A T T A G A A T G T T A C G A A A A A A C T A T G G T T G T G T C C A G T G E A T T T T	1540
1201	G C A C A G C C A T T T C C T T A A A A G G A A T T T A G A A A G C C A A A G T C A G A A A C C G G T G T C T G C T	1260
1541	G C A C A G C C A T T T C C T T A A A G G A A T T T A G A A A G C C A A A G T C A G A A A C C G G T G T C T G C T	1600
1261	C T A C T T T C C T T G A G C A A G C G T T T T A C C A G C T A T A C T T C C T T C A A G A C C A G T G A T G C T	1320
1601	C T A C T T T C C C T T G G A G C A A G C G T T T T A C C A G C T A T A C T T C C T T C A A G C C C A G T G A T G C T	1660
1321	G C T G A T G A A G G T A G A G A C A C G T C C A T T A A T G A G T C C A G A A T G C A C A G A T G A A T C C C T A	1380
1661	G C T G A T G A A G G T A G A G A C A C G T C C A T T A A T G A G T C C A G A A T G C A C A G A T G A A T C C C T A	1720
1381	C G A A G A G G T T G A T T G C A A A T C T G G C T G A G C A T A T C T A T T C A C T G C T A C A A G T C C T G T	1440
1721	C G A A G A G G T T G A T T G C A A A T C T G G C T G A G C A T A T T C T A T T C A C T G C T A C A A G T C C T G T	1780
1441	G C C A T T A T G C C A C A C A T T G T G G C T T G C C T C C T C T A C A G A C A C A G C A G G A A T T	1500
1781	G C C A T T A T G T C C A C A C A C A T T G T G G C T T G C C T C C T C T A C A G A C A C A G C A G G A A T T	1840
1501	G A T C T C T C C A C A T T G F T C G A A G A C T T C T T T G T G A T G A A A G A G A G A G T C T C G G C T C G T G A T	1560
1841	G A T C T C T C C A C A T T G F T C G A A G A C T T C T T T G T G A T G A A A G A G A G A G T C T C G G C T C G T G A T	1900
1561	T T T G A C C T G G G T T C T C A G G A A A T T C A G A A G A T G T A G T A A T G A T G A T G C C A T A C A G C T G C T G	1620
1901	T T T G A C C T G G G T T C T C A G G A A A T T C A G A A G A T G T A G T A A T G A T G A T G C C A T A C A G C T G C T G	1960
1621	G G A A A T T G T C A C A A T C A C C C A C A C T A G C A G A A C G A T G A G T T T T T A T C A C C C C C A G C	1680
1961	G G A A A T T G T C A C A A T C A C C C A C A C T A G C A G A A C G A T G A G T T T T T A T C A C C C C A G C	2020
1681	A C A A C T G T C C A T C A G T C T T C G A A C T C A A C T T C T A C A G A A A T G G G G T A C T T C A T G T C T T T	1740
2021	A C A A C T G T C C A T C A G T C T T C G A A C T C A A C T T C T A C A G A A A T G G G G T A C T T C A T G T C T T T	2080

Qy	1741	ATCATGGAGCCATCATAGCTTTGACGCTTTATGAGTTCTTGAAACAAGAGGGGACTGGGG	1800
Db	2081	ATCATGGAGGCCATCATAGCTTTGACGCTTTATGAGTTCTTGAAACAAGAGGGGACTGGGG	2140
Qy	1801	GGTCCCACTAGCACCCCACTAACTGATCAGCCAGGAGCAGCTGTCGCGAAAGCGGCC	1860
Db	2141	GGTCCCACTAGCACCCCACTAACTGATCAGCCAGGAGCAGCTGGTCGGAGGCGGCC	2200
Qy	1861	AGCCTGTGTACTCTTCTCCAAATGAAGGCACCATCTCACTGCCTTGCCAGACATTTTAC	1920
Db	2201	AGCCTGTGTACTCTTCTCCAAATGAAGGCACCATCTCACTGCCTTGCCAGACATTTTAC	2260
Qy	1921	CAAGTCTGCCATGAACACAGTAGGAAGTTTATCCAGTATGGCATTTCTTACAGTGGCAGAG	1980
Db	2261	CAAGTCTGCCATGAACACAGTAGGAAGTTTATCCAGTATGGCATTTCTTACAGTGGCAGAG	2320
Qy	1981	CACGATTGACAGGAAGATATCAGTCTCTAGTCTTGTCTGAGCAGCAGTGGGACAAGAAGCTT	2040
Db	2321	CACGATTGACAGGAAGATATCAGTCTCTAGTCTTGTCTGAGCAGCAGTGGGACAAGAAGCTT	2380
Qy	2041	CCAGAACCTTTGCTTGGAGAAGTGATGAAGAAGATGAAGACAGTGACTTTTGGGGAGGAA	2100
Db	2381	CCAGAACCTTTGCTTGGAGAAGTGATGAAGAAGATGAAGACAGTGACTTTTGGGGAGGAA	2440
Qy	2101	CACGGAGATTGCTACTGTAAGGTGAGCCATCCAAAGGAGCACCAGCAGTTTATCACCTTC	2160
Db	2441	CACGGAGATTGCTACTGTAAGGTGAGCCATCCAAAGGAGCACCAGCAGTTTATCACCTTC	2500
Qy	2161	TTACAGAGACTCTTGGGGCTTTTGTCTGGAGGCTACACGCTGCTGCCACTCTTTGTTTCAC	2220
Db	2501	TTACAGAGACTCTTGGGGCTTTTGTCTGGAGGCTACACGCTGCTGCCACTCTTTGTTTCAC	2560
Qy	2221	AACTTCAGTGGTCTGTTCCGAACCTCAGTATCTGCAAAAAGTTGSCAAAAATACCTAATA	2280
Db	2561	AACTTCAGTGGTCTGTTCCGAACCTCAGTATCTGCAAAAAGTTGSCAAAAATACCTAATA	2620
Qy	2281	ACCAGAAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCCACATATGCTGTGTGAAG	2340
Db	2621	ACCAGAAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCCACATATGCTGTGTGAAG	2680
Qy	2341	AATGCTGTGAAAATGTTTAAAGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTG	2400
Db	2681	AATGCTGTGAAAATGTTTAAAGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTG	2740
Qy	2401	TCTGTTTTTAAAGACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAAAACTTCTAGAA	2460
Db	2741	TCTGTTTTTAAAGACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAAAACTTCTAGAA	2800
Qy	2461	TATATTCTGAGTTTTGTGGTGCTG	2484
Db	2801	TATATTCTGAGTTTTGTGGTGCTG	2824

RESULT 4  
US-10-288-252-26  
; Sequence 26, Application US/10288252  
; Publication No. US20030134686A1  
; GENERAL INFORMATION:  
; APPLICANT: INQTE GENOMICS, INC.  
; APPLICANT: IAL, Preeti G.  
; APPLICANT: FANG, Y. Tom  
; APPLICANT: YUE, Henry  
; APPLICANT: BURFORD, Neil  
; APPLICANT: GANDHI, Aneena R.  
; APPLICANT: WARREN, Bridget A.  
; APPLICANT: YAO, Monique G.  
; APPLICANT: TRIOULEY, Catherine M.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: LEE, Ernestine A.  
; APPLICANT: HAFALIA, April J.A.  
; APPLICANT: LU, Yan  
; APPLICANT: GRIFFIN, Jennifer A.

```
; APPLICANT: SANJANWALA, Madhu S.
; APPLICANT: DING, Li
; TITLE OF INVENTION: TRANSFERASES
; FILE REFERENCE: PI-0241 USA
; CURRENT APPLICATION NUMBER: US/10/288,252
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT US01/30424
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/252,819
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/249,639
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/247,931
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/246,001
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/244,025
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/238,481
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,523
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 26
; LENGTH: 3210
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030143686A1 1517230CB1
US-10-288-252-26

Query Match      99.9%; Score 2480.8; DB 13; Length 3210;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2482; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ATGGATCAATCTGCACATGACCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 60
DB      338  ATGGATGAATCTGCATGACCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 397

QY      61  GAATACAGTGTGGTGCATGTAAGCACACAAGTGAGGAATGGGTGAGTGTGGCTTTAGA 120
DB      398  GAATACAGTGTGGTGCATGTAAGCACACAAGTGAGGAATGGGTGAGTGTGGCTTTAGA 457

QY      121  CCCACGCTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAATGAGTCGGAAGG 180
DB      458  CCCACCATCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAATGAGTCGGAAGG 517

QY      181  CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAATTTTCAAC 240
DB      518  CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAATTTTCAAC 577

QY      241  CCAGATATCCCGTCTTTGGGTTTGGGGAATGTTATTTATATCAATGAATTCACACAAGA 300
DB      578  CCAGATATCCCGTCTTTGGGTTTGGGGAATGTTATTTATATCAATGAATTCACACAAGA 637

QY      301  CACCGGGATGGCTTGCAAGACGCTTTCTTAGCTTTTATTTAAGAGGAGATGTG 360
DB      638  CACCGGGATGGCTTGCAAGACGCTTTCTTAGCTTTTATTTAAGAGGAGATGTG 697

QY      361  CATAAGGGCATGTTTGGCCCAATGTGACCTGAAATGTGCTGAACAGCAGTACAGTACAA 420
DB      698  CATAAGGGCATGTTTGGCCCAATGTGACCTGAAATGTGCTGAACAGCAGTACAGTACAA 757

QY      421  GAGCAATTCAGAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCACAGCAATCA 480
DB      758  GAGCAATTCAGAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCACAGCAATCA 817

QY      481  AAAGCCGTTAAACAAAGTGAAGAAAGAGCTTAAAGGATTTCTTCAAGAAATGGTTGCCACT 540
DB      818  AAAGCCGTTAAACAAAGTGAAGAAAGAGCTTAAAGGATTTCTTCAAGAAATGGTTGCCACT 877

541  GTCTCACCGCAATGATCAGACTGACTGGTGGGTGCTGCTAAACTGTTTCAACAGCTTC 600
878  GTCTCACCGCAATGATCAGACTGACTGGTGGGTGCTGCTAAACTGTTTCAACAGCTTC 937

601  TTTTGGAAACATTCAAATTCACAAAGGTCAACTTGAGATGGTTAAAGCTCAACTGAGAG 660
938  TTTTGGAAACATTCAAATTCACAAAGGTCAACTTGAGATGGTTAAAGCTCAACTGAGAG 997

661  AATTTGCGCGTCTGTTTCTTACAGTTCATAGATCCCATATTTGACTATCTGCTGCTCACT 720
998  AATTTGCGCGTCTGTTTCTTACAGTTCATAGATCCCATATTTGACTATCTGCTGCTCACT 1057

721  TTCAATTTCTTCTGCGCATAACTCAAGACCAATACATTCATTGCTTCAGGCAATATCTCAAC 780
1058  TTCAATTTCTTCTGCGCATAACTCAAGACCAATACATTCATTGCTTCAGGCAATATCTCAAC 1117

781  ATCCCAATCTTCAGTACCTTGATCCATTAAGCTTTGGGGGCTTCTTCAATACGAGAGGCTC 840
1118  ATCCCAATCTTCAGTACCTTGATCCATTAAGCTTTGGGGGCTTCTTCAATACGAGAGGCTC 1177

841  GATGAACACACAGATGGACGGAAGAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900
1178  GATGAACACACAGATGGACGGAAGAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1237

901  GTTGAATTAATTCGACAGCAGCAATTTCTTGGAGATCTTCTCGAAGGCAACAGTTCAGG 960
1238  GTTGAATTAATTCGACAGCAGCAATTTCTTGGAGATCTTCTCGAAGGCAACAGTTCAGG 1297

961  AGTGGAAAAACCTCTTGTGCTCGGAGAGCTTTTGTGTCAGTTGCTGAGTGTGATGCTGTCT 1020
1298  AGTGGAAAAACCTCTTGTGCTCGGAGAGCTTTTGTGTCAGTTGCTGAGTGTGATGCTGTCT 1357

1021  ACCAATGTCTATCCAGACATCTTGATATACCTGTTGGAATCTCTATGATCGCATATC 1080
1358  ACCAATGTCTATCCAGACATCTTGATATACCTGTTGGAATCTCTATGATCGCATATC 1417

1081  GAAGGTCACTACAATGGTGAAACAACTGGGCAACCTTAAGAAGAAATGAGAGCTGTGGAGT 1140
1418  GAAGGTCACTACAATGGTGAAACAACTGGGCAACCTTAAGAAGAAATGAGAGCTGTGGAGT 1477

1141  GTAGCAAGAGTGTGTTATTAAGATGTTACGAAAAAACTATGTTGTTGCTCGAGTGGATTTT 1200
1478  GTAGCAAGAGTGTGTTATTAAGATGTTACGAAAAAACTATGTTGTTGCTCGAGTGGATTTT 1537

1201  GCACAGCAATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGCTGCT 1260
1538  GCACAGCAATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGCTGCT 1597

1261  CTACTTTTCCCTCGAGCAAGCTTGTATCCAGCTATACCTTCTTCAAGACCCAGTGTGCT 1320
1598  CTACTTTTCCCTCGAGCAAGCTTGTATCCAGCTATACCTTCTTCAAGACCCAGTGTGCT 1657

1321  GCTGATGAAGGTGAGAGACAGCTCCATTAATAGTCCAGAAATGCAACAGATGAATCCCTTA 1380
1658  GCTGATGAAGGTGAGAGACAGCTCCATTAATAGTCCAGAAATGCAACAGATGAATCCCTTA 1717

1381  CGAAGAGGTGCTATGTCGAATCTGGCTGAGCATATTTATTCACCTGCTAGCAAGTCTGCT 1440
1718  CGAAGAGGTGCTATGTCGAATCTGGCTGAGCATATTTATTCACCTGCTAGCAAGTCTGCT 1777

1441  GCATTTATGTCACACACATCTGCTTGGCTTCCCTGCTCTCTACAGACACAGGAGGAAAT 1500
1778  GCATTTATGTCACACACATCTGCTTGGCTTCCCTGCTCTCTACAGACACAGGAGGAAAT 1837

1501  GATCTCTCCATTTGGTGAAGACCTTCTTTGATGAAGAGGAAAGTCTCGCTGCTGAT 1560
1838  GATCTCTCCATTTGGTGAAGACCTTCTTTGATGAAGAGGAAAGTCTCGCTGCTGAT 1897

1561  TTTTGACCTGGGCTTCTCAGGAAATTCAGAGATGTAGTAATCATGCCCATACAGCTGCTG 1620
1898  TTTTGACCTGGGCTTCTCAGGAAATTCAGAGATGTAGTAATCATGCCCATACAGCTGCTG 1957

1621  GGAATTTGTGTCACAATCACCCCACTAGCAGGAACGATGAGTTTTTTTATCACCACAGC 1680
```

Db 1958 GAAATTTGTGCACAAATCACCACACTAGCAGGAACGATGATTTTTTATCACCCCGACG 2017  
Qy 1681 ACAACTGTCCCATCAGTCTTGGAACCTCAACTTCTACAGCAATGGGGTACTTTCATGTCTTT 1740  
Db 2018 ACAACTGTCCCATCAGTCTTGGAACCTCAACTTCTACAGCAATGGGGTACTTTCATGTCTTT 2077  
Qy 1741 ATCATGGAGGCATCATAGTCTGACGCTTTATGCAAGTCTTGAACAAGAGGGGACTGGG 1800  
Db 2078 ATCATGGAGGCATCATAGTCTGACGCTTTATGCAAGTCTTGAACAAGAGGGGACTGGG 2137  
Qy 1801 GGTCCCACTAGCACCCCACTAACCCTGATCAGCAGGAGCAGCTGGTGGGAAGCGGCC 1860  
Db 2138 GGTCCCACTAGCACCCCACTAACCCTGATCAGCAGGAGCAGCTGGTGGGAAGCGGCC 2197  
Qy 1861 AGCTGTGCTACTTCTCTCCAATGAAGCGCACCACTCTCACTGCGCTTGCGACACATTTTAC 1920  
Db 2198 AGCTGTGCTACTTCTCTCCAATGAAGCGCACCACTCTCACTGCGCTTGCGACACATTTTAC 2257  
Qy 1921 CAACTGTCCCATGAACAAGTAGGAAAGTTTATCCAGTATGGGATTTTACAGTGGCAGAG 1980  
Db 2258 CAACTGTCCCATGAACAAGTAGGAAAGTTTATCCAGTATGGGATTTTACAGTGGCAGAG 2317  
Qy 1981 CACGATGACAGGAGATATCAGTCTAGTCTTGTGAGCAGCAGTGGGCAAGAGCTT 2040  
Db 2318 CACGATGACAGGAGATATCAGTCTAGTCTTGTGAGCAGCAGTGGGCAAGAGCTT 2377  
Qy 2041 CCAGAACCTTGTCTTGTGAGAGATGATGAAGAAGATGAAGACAGTGAATTTGGGGAGGAA 2100  
Db 2378 CCTGAACCTTGTCTTGTGAGAGATGATGAAGAAGATGAAGACAGTGAATTTGGGGAGGAA 2437  
Qy 2101 CAGCGAGATGCTACCTGAAGGTGAGCCAAATCCAGGAGCACCAGCAGTATCACCTTC 2160  
Db 2438 CAGCGAGATGCTACCTGAAGGTGAGCCAAATCCAGGAGCACCAGCAGTATCACCTTC 2497  
Qy 2161 TTACAGACACTCCTTGGGCGCTTGTGAGGCGCTACAGCTCTGCTGCCATCTTTGTTCC 2220  
Db 2498 TTACAGACACTCCTTGGGCGCTTGTGAGGCGCTACAGCTCTGCTGCCATCTTTGTTCC 2557  
Qy 2221 AACTTCAGTGTCTGTTCAGAACCTCAGTATCTGCAAAAGTTGCAAAATACCTAATA 2280  
Db 2558 AACTTCAGTGTCTGTTCAGAACCTCAGTATCTGCAAAAGTTGCAAAATACCTAATA 2617  
Qy 2281 ACCAGAACAGAAAGAAATTTGCAAGTATATGCTGAGAGTGCACATATTTGTTGAAG 2340  
Db 2618 ACCAGAACAGAAAGAAATTTGCAAGTATATGCTGAGAGTGCACATATTTGTTGAAG 2677  
Qy 2341 AATGCTGAAATGTTTAAAGATATTTGGGGTTTCAAGGAGACCAAAACAAAGAGAGTG 2400  
Db 2678 AATGCTGAAATGTTTAAAGATATTTGGGGTTTCAAGGAGACCAAAACAAAGAGAGTG 2737  
Qy 2401 TCTGTTTTAGAACCTGAGCAGCACTTTTCTACCTCAATGCAACCGCACAAAACCTTCTAGAA 2460  
Db 2738 TCTGTTTTAGAACCTGAGCAGCACTTTTCTACCTCAATGCAACCGCACAAAACCTTCTAGAA 2797  
Qy 2461 TATATCTGAGTTTGTGGTGCTG 2484  
Db 2798 TATATCTGAGTTTGTGGTGCTG 2821

RESULT 5

US-09-917-800A-1648

; Sequence 1648, Application US/09917800A

; Patent No. US20020119462A1

; GENERAL INFORMATION:

; APPLICANT: Mendrick, Donna

; APPLICANT: Porter, Mark

; APPLICANT: Johnson, Kory

; APPLICANT: Castle, Arthur

; APPLICANT: Elashoff, Michael

; APPLICANT: Gene Logic, Inc.

; TITLE OF INVENTION: Molecular Toxicology Modeling

; FILE REFERENCE: 44921-5038-US

; CURRENT APPLICATION NUMBER: US/09/917,800A

; CURRENT FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: US 60/222,040

; PRIOR FILING DATE: 2000-07-31

; PRIOR APPLICATION NUMBER: US 60/222,880

; PRIOR FILING DATE: 2000-11-02

; PRIOR APPLICATION NUMBER: US 60/290,029

; PRIOR FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: US 60/290,645

; PRIOR FILING DATE: 2001-05-15

; PRIOR APPLICATION NUMBER: US 60/292,336

; PRIOR FILING DATE: 2001-05-22

; PRIOR APPLICATION NUMBER: US 60/295,798

; PRIOR FILING DATE: 2001-06-06

; PRIOR APPLICATION NUMBER: US 60/297,457

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/298,884

; PRIOR FILING DATE: 2001-06-19

; PRIOR APPLICATION NUMBER: US 60/303,459

; PRIOR FILING DATE: 2001-07-09

; NUMBER OF SEQ ID NOS: 1740

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1648

; LENGTH: 2646

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM\_017274

US-09-917-800A-1648

Query Match

Best Local Similarity 75.7%; Score 1879.2; DB 10; Length 2646;

Matches 2106; Conservative 0; Mismatches 378; Indels 0; Gaps 0;

Qy 1 ATGATCAATCTGACATGACCCCTGGTACATAGATGTTCTTATCTGCCACATTCATCA 60  
Db 160 ATGAGGAGTCTTCAGTGACAAATGGCAACAATAGACGTTCTTATCTGCCAATTCATCA 219  
Qy 61 GAATACAGTGTGGTCCATGTAAAGCACACAAAGTGGGCTGAGTGGCTTTAGA 120  
Db 220 GAATACAGCTTGGCCCATGTAAACACCAATGAGGACTGGGTGACTGGCTTCAAA 279  
Qy 121 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCCTAATGAGTCGGAAGG 180  
Db 280 CCTACCTCTTCAGATCGCAACCGCTGAATGGAAGAGAGCCCTCATGAGCGGAAGG 339  
Qy 181 CCATTTGTGGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAATTTTCAAC 240  
Db 340 CCTTCGTGGAGAGTGTGCTATTTCATGCAACGCTCAGAGCTGGGAAAGGTTTTTCAAC 399  
Qy 241 CCAGTATCCGCTCTTTGGGTTTGGCGAATGTTATTATATCAATGAACTCACACAAGA 300  
Db 400 CCAGTATCCATCTCTGGGTTTGGCGAATGTTATTATATCAATGAACTCACACAAGG 459  
Qy 301 CACCGCGATGGCTTGAAGACGCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTG 360  
Db 460 CACCGAGATGGCTGGCAAGACGCTTTCTTACATCTTTTGTTCAGAGCGGATGTC 519  
Qy 361 CATAGGGCATGTTTGGCCACCAATGTGCTGAAATGTGCTGAACAGCAGTAGAGTCAA 420  
Db 520 CACAAGGGCATGTTTGGCCACCAATGTGCTGAAATGTGCTGAACAGCAGTAGAGTCAA 579  
Qy 421 GAGCAATTCAGAGTGGCTGTGAATTAACCCCTGATGTTCTTCTGCCACGACCAATCA 480  
Db 580 GAGCAATTCAGAGTGGCTGTGAATTAACCCCTGATGTTCTTCTGCCACGACGATCC 639  
Qy 481 AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTTTTCAAGAAATGTTTGCCTACT 540  
Db 640 AAAGCCATCCAGAAAGTGAAGAAAGAAAGCCAGGAGATCTCCAGGAAATGTTGCTACA 699  
Qy 541 GTCTCACCGCAATGATCAGACTGAGTGGGTGCTGCTTAAACCTGTTCAACAGCTTC 600  
Db 700 GTCTCCCCGGGATGATCAGCTGACTGGGTGTTTACTTAAAGCTCTTCAACAGCTTC 759

Qy	601	TTTTGGAA	CAATTC	AAATTT	CA	AAAGGTC	AATCTTG	AGATGG	TTAAAGCTG	CAACTG	AGACG	660
Db	760	TTCTGGAA	CAATTC	AGATTT	CA	AAAGGTC	AACTTG	AGATGG	TGAAGCTG	CAACTG	AGACG	819
Qy	661	AAATTTGG	CGGCTT	CTGGTTT	CT	ACCAAGT	TCATAG	ATCCCAT	TTGACTAT	CTGCTG	CTCACT	720
Db	820	AAATCTGG	CGCTT	TTGTTT	CT	CGCGGTG	CAAGAT	TC	CCACATCG	ACTTAC	CTGCTC	879
Qy	721	TTCAATCT	CTTTG	CCATAA	CAATCA	AAAGCACC	ATACAT	TGCTT	CAGGCA	ATAATCT	CAAC	780
Db	880	TTCAATCT	CTTTG	CCACAACA	CAATCA	AAAGCTCC	ATACAT	CGCTT	CGGCA	CACACCT	CAAC	939
Qy	781	ATCCCAAT	CTTCA	GTACCT	TTAGT	CCATPA	AGCTT	GGGGCTT	CTTTC	CATACGA	CGAGCTC	840
Db	940	ATCCCAAT	CTTCA	GTACCT	TTAGT	TTCA	AAAGCTT	TGGGGCTT	TTTTTC	ATAAGA	CGAGGCTT	999
Qy	841	GATGAACA	CCAGATG	CAACG	GAAGATG	TCTCT	ATAGAGCTT	TGCTTC	CAATGG	GCATATA	900	
Db	1000	GACGAACA	CTCAG	ATGGA	CGCAAG	ACATTT	CTGTAC	AGAGCGT	TGCTTC	CAATGG	GCATATA	1059
Qy	901	GTTGAA	TTACTT	CGACAG	CACTT	CTTGAG	ATCTT	CTGAG	ATCTT	CTGGAAG	GCACACG	960
Db	1060	GTTGAA	CTCCT	CGACAG	CACTT	CTTGAG	ATCTT	CTGAG	ATCTT	CTGGAAG	GCACACG	1119
Qy	961	AGTGGAAA	ACCTT	CTTGCT	CGGCA	AGCAT	TTTTT	TGTCA	GTTG	GTGTAG	ATATCT	1020
Db	1120	AGTGGCA	AGACCT	CTCTG	CGCGG	CGGGCT	CTCTG	TCA	GTTG	GTGTAG	ATATCT	1179
Qy	1021	ACCAATG	TCATCC	CAGAC	ATCTT	GTATAT	ACCTT	GTGGA	ATCTT	CTATGAT	CGATATC	1080
Db	1180	TCCAACA	ACCATCC	CTGAC	ATCTT	GTCAT	CTCTG	TGGGCA	TCTCG	TATGAT	CGGATATC	1239
Qy	1081	GAAGGT	CACTACA	ATG	TGTGA	CAACT	CGGCA	AACTT	AGAAGA	ATAGAG	CGCTGTG	1140
Db	1240	GAAGGT	CACTACA	ATG	TGTGA	CAACT	CGGCA	AACTT	AGAAGA	ATAGAG	CGCTGTG	1299
Qy	1141	GTAGCA	AGAGGT	GTAT	TAGAA	TGTTA	CGAAAA	AACTAT	TGTTG	TGTCT	CGAGTGG	1200
Db	1300	GTGCA	AGAGCG	GTAT	CAGAA	TCTCG	GAAAAA	AACTAT	CGGCTAT	GTCTCG	AGTGG	1359
Qy	1201	GCACAG	CCATTT	CTCTT	AAAGGA	ATAT	TAGA	AGCCAA	AGTC	AGAAC	CGGTGT	1260
Db	1360	GCACAG	CCATTT	CTCT								

RESIST. T. 6

```

RES001 6
US-09-833-381-1961
; Sequence 1961, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1961
; LENGTH: 612

```



```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1961

Query Match      23.4%; Score 581.6; DB 10; Length 612;
Best Local Similarity 99.0%; Pred. No. 9.1e-175;
Matches 606; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy 1061 TCTCCTATGATCGCATTCGAAGTCTACTCAATGTGTGAACAACTGGGCAAACTTAAGA 1120
Db 1 TCCCGTATGATCGCATTCGAAGTCTACTCAATGTGTGAACAACTGGGCAAACTTAAGA 60

Qy 1121 AGAATGAGACCTGTGAGTGTAGCAAGAGTGTATTAGAAATGTTAGCAAAAATATG 1180
Db 61 AGAATGAGACCTGTGAGTGTAGCAAGAGTGTATTAGAAATGTTAGCAAAAATATG 120

Qy 1181 GTTGTGTCGAGTGTGATTTTCACAGCCATTTCTTAAAGGAATATTAGAAAGCCAAA 1240
Db 121 GTTGTGTCGAGTGTGATTTTCACAGCCATTTCTTAAAGGAATATTAGAAAGCCAAA 180

Qy 1241 GTCAGAAA - CCGGTGTCGTCTACTTTCCCTGGAGCAAGCGTTGTACCAAGCTATACTT 1299
Db 181 GTCAGAAAACCGGTGTCGTCTACTTTCCCTGGAGCAAGCGTTGTACCAAGCTATACTT 240

Qy 1300 CCTTCAGACCCAGTGTGCTGTGATGAAGGTAGAGACAGTCCATTATGATGCCAGA 1359
Db 241 CCTTCAGACCCAGTGTGCTGTGATGAAGGTAGAGACAGTCCATTATGATGCCAGA 300

Qy 1360 AATCAACAGATGAATCCCTACGAAGGAGTTGATGCAATCTGGCTGAGCATATCTA 1419
Db 301 AATCAACAGATGAATCCCTACGAAGGAGTTGATGCAATCTGGCTGAGCATATCTA 360

Qy 1420 TTCACTGCTAGCAAGTCTGTGCCATTATGCCACACACATTTGGCTTGCCTCTCTC 1479
Db 361 TTCACTGCTAGCAAGTCTGTGCCATTATGCCACACACATTTGGCTTGCCTCTCTC 420

Qy 1480 TACAGACAGGAGGGAATGATCTCTCCACATTTGTCGAAGACTCTTTTGTGATGAAA 1539
Db 421 TACAGACAGGAGGGAATGATCTCTCCACATTTGTCGAAGACTCTTTTGTGATGAAA 480

Qy 1540 GAGGAAGTCCTGGCTCGTGAATTTGACCTGGGTTCTCAGGAATTCAGAGATGT-AGT 1598
Db 481 GGGGAAGTCCTGGCTCGTGAATTTGACCTGGGTTCTCAGGAATTCAGAGATGTAAGT 540

Qy 1599 AATCATGCCATACAGCTGTCTGGGAAATTTGTCAATACACCCACACTAGCAGGAACGA 1658
Db 541 AATCATGCCATACAGCTGTCTGGGAAATTTGTCAATACACCCACACTAGCAGGAACCA 600

Qy 1659 TGAGTTTTTTAT 1670
Db 601 TGAGTTTTTTAT 612
```

```
RESULT 7
US-09-918-995-21644
; Sequence 21644, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21644
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: (1)...(482)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-21644

Query Match      17.6%; Score 437; DB 11; Length 482;
Best Local Similarity 100.0%; Pred. No. 1.6e-128;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 264 GCGGAATGTTATTATATCAATGAATCAACAGACACCGCGGATGGCTTGAAGACG 323
Db 46 GCGGAATGTTATTATATCAATGAATCAACAGACACCGCGGATGGCTTGAAGACG 105

Qy 324 CTTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTGCATAAGGGCATGTTTGCACCAA 383
Db 106 CTTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTGCATAAGGGCATGTTTGCACCAA 165

Qy 384 TGTGACTGAAATATGCTGAAACAGCAGTAGAGTACAAGAGCAATTTGAGAAGTGGTGC 443
Db 166 TGTGACTGAAATATGCTGAAACAGCAGTAGAGTACAAGAGCAATTTGAGAAGTGGTGC 225

Qy 444 TGAATTAACCCCTGATGGTTCTGCCAGCAGCAATCAAAAGCCGTTAAACAAAGTGAATA 503
Db 226 TGAATTAACCCCTGATGGTTCTGCCAGCAGCAATCAAAAGCCGTTAAACAAAGTGAATA 285

Qy 504 GAAAGCTAAAGAGATTTCTTCAAGAAATGGTTGCCACTGTCTCACCGGCAATGATCAGACT 563
Db 286 GAAAGCTAAAGAGATTTCTTCAAGAAATGGTTGCCACTGTCTCACCGGCAATGATCAGACT 345

Qy 564 GACTGGTGGTGTCTGCTTAAACCTGTTCAACAGCTTTTGGAAACATTTCAAAATTCACAA 623
Db 346 GACTGGTGGTGTCTGCTTAAACCTGTTCAACAGCTTTTGGAAACATTTCAAAATTCACAA 405

Qy 624 AGGTCAACTTGAGATGTTAAAGCTCAACTGACAGCAATTTGCGCTTCTGTTTCTACC 683
Db 406 AGGTCAACTTGAGATGTTAAAGCTCAACTGACAGCAATTTGCGCTTCTGTTTCTACC 465

Qy 684 AGTTTCATAGATCCCAT 700
Db 466 AGTTTCATAGATCCCAT 482

RESULT 8
US-09-918-995-7822
; Sequence 7822, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7822
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-7822

Query Match      15.5%; Score 385.4; DB 11; Length 399;
Best Local Similarity 99.5%; Pred. No. 4.5e-112;
Matches 397; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 290 CTCACACAAGACACCGCGGATGGCTTGAAGAGCCCTTTCTTACGTTCTTTTATCAAG 349
Db 1 CTCACACAAGACACCGCGGATGGCTTGAAGAGCCCTTTCTTACGTTCTTTTATCAAG 60

Qy 350 AGCGAGATGTGCATAAGGGGATGTTTCCCAACCAATGTGACTGAAAAATGTGC-TGAACAGC 408
```



Db 61 AGCAGATGTGCATAGAGGCGATGTTTCCACCAATGTGACTGAAATGTGCTTTGAACAGC 120  
QY 409 AGTAGAGTACAAGAGCAATTGACAGAGTGGCTGCTGAATTAACCCCTGATGTTCTGCC 468  
Db 121 AGTAGAGTACAAGAGCAATTGACAGAGTGGCTGCTGAATTAACCCCTGATGTTCTGCC 180  
QY 469 CAGCAGCAATCAAAAGCCGTTTACAAGTGAAGAAAGAAAGCTAAAGAGGATTTCTTCAAGAA 528  
Db 181 CACCAGCAATCAAAAGCCGTTTACAAGTGAAGAAAGAAAGCTAAAGAGGATTTCTTCAAGAA 240  
QY 529 ATGGTTGCCACTGCTCTCACCGCAATGATCAGACTGACTGGGTGGTGTCTGCTAAAGCTG 588  
Db 241 ATGGTTGCCACTGCTCTCACCGCAATGATCAGACTGACTGGGTGGTGTCTGCTAAAGCTG 300  
QY 589 TTCAACAGCTCTTTTGGACATTCATAATTCACAAAGGTCACCTTCAGATGTTTAAAGCT 648  
Db 301 TTCAACAGCTCTTTTGGACATTCATAATTCACAAAGGTCACCTTCAGATGTTTAAAGCT 360  
QY 649 GCAACTGAGACGAATTTGCCGCTTCTGTTTCTACCAAGTT 687  
Db 361 GCAACTGAGACGAATTTGCCGCTTCTGTTTCTACCAAGTT 399

## RESULT 9

US-09-833-381-1964  
; Sequence 1964, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1964  
; LENGTH: 440  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(440)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1964

Query Match 14.0%; Score 348.4; DB 10; Length 440;  
Best Local Similarity 98.3%; Pred. No. 3.5e-100;  
Matches 349; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 87 CACAAGTAGGAATGGGGTGAAGTGGCTTTAGACCCACCGTCTTCAGATCTGCAACTTT 146  
Db 15 CCCAAGTGAGGAATGGGGTGAAGTGGCTTTAGACCCACCGTCTTCAGATCTGCAACTTT 74  
QY 147 ABAATGGAAGAAAGCCTAATAGTCGGAAGGCCATTTGTTGGAAGATGTTGTATCTC 206  
Db 75 ABAATGGAAGAAAGCCTAATAGTCGGAAGGCCATTTGTTGGAAGATGTTGTATCTC 134  
QY 207 CTGCACCTCCCGAGAGCTGGGACAAAATTTTCAACCCCGATATCCCGTCTTTGGGTTGGC 266  
Db 135 CTGCACCTCCCGAGAGCTGGGACAAAATTTTCAACCCCGATATCCCGTCTTTGGGTTGGC 194  
QY 267 GAATGTTATTTATCATGAAACTACACAAGACACCGCGGATGGCTTGCAGAGCGCT 326  
Db 195 GAATGTTATTTATCATGAAACTACACAAGACACCGCGGATGGCTTGCAGAGCGCT 254  
QY 327 TTCTTAGGTTCTTTTATTTCAAGCGAGATGTCATAAGGCGATGTTTGCACCAATGT 386  
Db 255 TTCTTAGGTTCTTTTATTTCAAGCGAGATGTCATAAGGCGATGTTTGCACCAATGT 314  
QY 387 GACTGAAATGTGCTGAACAGCAGTAGAGTACAAGAGGCAATTTGCAGAGTGGCT 441

Db 315 GACTGAAATGTGCTTAAACNGCANTAGAGTACAAGAGCAATTCGNNAAAGTGGCT 369  
RESULT 10  
US-09-833-381-1959  
; Sequence 1959, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1959  
; LENGTH: 396  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(396)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1959

Query Match 14.0%; Score 346.6; DB 10; Length 396;

Best Local Similarity 95.9%; Pred. No. 1.2e-99;  
Matches 375; Conservative 0; Mismatches 13; Indels 3; Gaps 2;  
QY 1838 AGCAGCTGGTGGGAAGCGCGCAGCTGTGTACCTTCTTCCCAATGAAGCACCATCT 1897  
Db 1 ACAGCTNGTGGGAAGCGCGCAGCTGTGTACCTTCTTCCCAATGAAGCACCATCT 60  
QY 1898 CACTGCTTGGCAGACATTTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTATCCAGT 1957  
Db 61 CACTGCTTGGCAGACATTTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTATCCAGT 120  
QY 1958 ATGGCATTTTACAGTGGCAGACGATGACCAGGAAGATATCAGTCTAGCTTGTCTG 2017  
Db 121 ATGGCATTTTACAGTGGCAGACGATGACCAGGAAGATATCAGTCTAGCTTGTCTG 180  
QY 2018 AGCAGCTGGGACAAAGCTTCCAGAACCTTTCTTGGGAAGTGTATGAAGAAGATG 2077  
Db 181 AGCAGCTGGGACAAAGCTTCTTGAACTTTCTTGGGAAGTGTATGAAGAAGATG 240  
QY 2078 AAGACAGTGAATTTGGGAGGAACAGCGAGATTGCTACTGAA--GGTGAGCCAAATCCAA 2135  
Db 241 AAGACAGTGAATTTGGGAGGAACAGCGAGATTGCTACTGAAAGGTGAGCCCAATCCAA 300  
QY 2136 GGAGCACCAGCAGTTTATCACCCTTTTACAGAGACTCTTGGGCTTTGCT--GGAGGCT 2194  
Db 301 GGAGCACCAGCAGTTTAAACAACCTTCTTACAGAGTCTCTTGGGCTTTGCTGGAGGCT 360  
QY 2195 ACAGCTCTGCTGCCATCTTTGTTTCAACTT 2225  
Db 361 ACAGCTCTGCTGCCATCTTTGTTTCAACTT 391

## RESULT 11

US-09-960-352-7384  
; Sequence 7384, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Ningbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352

```

; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7384
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 32-LIB2809-019-Q1-E1-H11
US-09-960-352-7384

Query Match      13.8%; Score 342.2; DB 10; Length 429;
Best Local Similarity 90.2%; Pred. No. 3.3e-98;
Matches 378; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

QY 1330 GGTAGAGACAGTCCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAGGAGG 1389
DB 14 GATACAGACAAAGTCCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCGCCGA---AGA 70

QY 1390 TTGATTGCAAAATCGGCTGAGCATATCTATTCACTGCTAGCAAGTCTGTGCCATTATG 1449
DB 71 CTGATTGACATCTGGCTGAGCACAATCTCTTCACTGCTAGCAAGTCTGTGCTATTATG 130

QY 1450 TCCACACATGTTGGCTTGGCTGCTCTCTACAGACAGGAGGAAATGATCTCTCC 1509
DB 131 TCCAGCAGATCTGACCTGTCTCTCTCTACAGACAGGAGGAAATGAGGCTCTCC 190

QY 1510 ACATTGGTCGAGACTTCTTTGTGATGAAAGGAAGTCTTGCTCGTGAATTTTGACCTG 1569
DB 191 ACATTGGTGAAGACTTCTTTGTGATGAAAGGAAGTCTTGCTCGTGAATTTTGACTTG 250

QY 1570 GGGTTCTCAGGAAATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1629
DB 251 GGGTTCTCAGGAAATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 310

QY 1630 GTCACAAATCACCACACTAGCAGGAAATGATGATGATGATGATGATGATGATGATGATG 1689
DB 311 ATCAATATCACCACACTAGCAGGAAATGATGATGATGATGATGATGATGATGATGATGATG 370

QY 1690 CCATCAGTCTTGAAGTCTTCTACAGCAATGAGGATGATGATGATGATGATGATGATGATG 1748
DB 371 GCATCAGTCTTGAAGTCTTCTACAGCAATGAGGATGATGATGATGATGATGATGATGATGATG 429

RESULT 12
US-09-960-352-8821
; Sequence 8821, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 8821
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 38-LIB2809-028-Q1-E1-B6
US-09-960-352-8821

Query Match      13.6%; Score 337.6; DB 10; Length 407;
Best Local Similarity 90.9%; Pred. No. 9.4e-97;
Matches 371; Conservative 0; Mismatches 34; Indels 3; Gaps 1;

QY 1330 GGTAGAGACAGTCCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAGGAGG 1389
DB 3 GATACAGACAAAGTCCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCGCCGA---AGA 59

QY 1390 TTGATTGCAAAATCTGGCTGAGCATATCTATTCACTGCTAGCAAGTCTCTGTGCCATTATG 1449
```

```

DB 60 CTGATTGCACATCTGGCTGAGCACAATCTCTTCACTGCTAGCAAGTCTGTGCTATTATG 119
QY 1450 TCCACACACATTGTTGGCTTGGCTGCTCTCTACAGACAGGAGGAAATGATGATCTCTCC 1509
DB 120 TCCACGACATCTGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 179
QY 1510 ACATTGGTCGAAAGACTTCTTTGTGATGAAAGGAAGTCTTGCTCGTGAATTTTGACCTG 1569
DB 180 ACATTGGTGAAGACTTCTTTGTGATGAAAGGAAGTCTTGCTCGTGAATTTTGACTTG 239
QY 1570 GGGTTCTCAGGAAATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1629
DB 240 GGGTTCTCAGGAAATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 299
QY 1630 GTCAAAATCACCACACTAGCAGGAAATGATGATGATGATGATGATGATGATGATGATG 1689
DB 300 ATCAATATCACCACACTAGCAGGAAATGATGATGATGATGATGATGATGATGATGATG 359
QY 1690 CCATCAGTCTTGAAGTCTTCTACAGCAATGAGGATGATGATGATGATGATGATGATGATG 1737
DB 360 CCATCAGTCTTGAAGTCTTCTACAGCAATGAGGATGATGATGATGATGATGATGATGATG 407

RESULT 13
US-09-833-381-1958/c
; Sequence 1958, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1958
; LENGTH: 868
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(868)
; OTHER INFORMATION: n = A, T, C or G
US-09-833-381-1958

Query Match      13.4%; Score 331.8; DB 10; Length 868;
Best Local Similarity 99.1%; Pred. No. 1.2e-94;
Matches 333; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2149 TTTATCACCCTTTTACAGAGACTCTCTTGGGCTTTTGTGGAGGCTTACAGCTCTGTGCTCC 2208
DB 865 TTTATCACCCTTTTACAGAGACTCTCTTGGGCTTTTGTGGAGGCTTACAGCTCTGTGCTCC 806
QY 2209 ATCTTTCTTACACACTTTCAGTGGTCTTTCAGAAATGATGATGATGATGATGATGATGATG 2268
DB 805 ATCTTTCTTACACACTTTCAGTGGTCTTTCAGAAATGATGATGATGATGATGATGATGATG 746
QY 2269 AAATACCTTAATACACAGAAAGAAATGTTTTCAGTATATGCTGAGAGTCCACATAT 2328
DB 745 AAATACCTTAATACACAGAAAGAAATGTTTTCAGTATATGCTGAGAGTCCACATAT 686
QY 2329 TGTCTTGTGAAGATCTGTGAAATGTTTAAAGATATTTGGGTTTTCAGAGAGACCAAA 2388
DB 685 TGTCTTGTGAAGATCTGTGAAATGTTTAAAGATATTTGGGTTTTCAGAGAGACCAAA 626
QY 2389 CAAAAGAGTGTCTGTTTGAAGTCTTTCAGAGAGTCTTTCACCTCAATGCAACCGCAA 2448
DB 625 CAAAAGAGTGTCTGTTTGAAGTCTTTCAGAGAGTCTTTCACCTCAATGCAACCGCAA 566
QY 2449 AAATCTTCTAGATATATTTCTGAGTCTTTCAGTCTTTCAGTCTGCTG 2484
```

Db  
565 AAACCTCTAGAATAATCTGAGTTTGTGGTGCIG 530

## RESULT 14

US-09-960-352-742/C  
 / Sequence 742, Application US/09960352  
 / Patent No. US20020137139A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Warren, Wesley C.  
 / APPLICANT: Tao, Nengbing  
 / APPLICANT: Byatt, John C.  
 / APPLICANT: Mathalagan, Nagappan  
 / TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
 / TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION  
 / FILE REFERENCE: 16511.006/37-21(10298)C  
 / CURRENT APPLICATION NUMBER: US/09/960,352  
 / CURRENT FILING DATE: 2001-09-24  
 / NUMBER OF SEQ ID NOS: 15112  
 / SEQ ID NO 742  
 / LENGTH: 418  
 / TYPE: DNA  
 / ORGANISM: Bos taurus  
 / FEATURE:  
 / NAME/KEY: unsure  
 / LOCATION: (395)  
 / OTHER INFORMATION: unsure at all n locations  
 / OTHER INFORMATION: Clone ID: 04-LIB2809-001-Q1-E1-A11  
 US-09-960-352-742

Query Match	12.2%	Score 302;	DB 10;	Length 418;
Best Local Similarity	84.1%	Pred. No. 2.4e-85;		
Matches 354;	Conservative 0;	Mismatches 61;	Indels 6;	Gaps 1;
QY	1772	ATCAGTTCGACAGAGAGGGGACTGGGGGTCCCACATAGCACACCCACCTAAACCTGATCA	1831	
DB	418	ATCAGATTCTGAAGAAGAGGGGNCAGGAGGACC-----CGGTCTCCCAGCTTGGTCA	365	
QY	1832	GCACGAGCAGCTGGTGCGGAAGCGGCCAGCCCTGTGCTACCTTCTCTCCAATGAAGGCA	1891	
DB	364	GCACGAGCAGCTGGTGCAACAGGCTGCCAGCTGTCTATCTGCTCTCCAATGAAGGCA	305	
QY	1892	CCATCTCACTGCCTTGCACAGCATTTTACCAAGTCTGCCATGAACAGTAGGAAAGTTTA	1951	
DB	304	CCATCTCTCTCCCTGCCAGACCTTTTACCAAAATTGCCATGAAACAGTGGCGCGTTTA	245	
QY	1952	TCCAGTATGGCATTTCTTACAGTGGCAGAGCAGATGACCCAGGAAAGATATCAGTCTTAGTC	2011	
DB	244	TCCAGTATTGCCAATTTTATATAGTTGCCGAGCAAGATAATCAGGAGAGATATCAGTTCCTTGT	185	
QY	2012	TTGCTGAGCAGCAGTGGGACAAAGAGCTTCCAGAACCTTTGTCTTGAGAGAGTGATCAAG	2071	
DB	184	CTTCCGAGCAGCAGTTGGACAAGAAGGTTCCGGAACTTTGTCTTGAGAGAGTGATGAAG	125	
QY	2072	AAGATGAAGACAGTGACTTTTGGGAGGAAACAGCGAGATTGCTACTCGAAGTGAGCCAAT	2131	
DB	124	AAGATGAAGACAGGATTTTGGTGAGAGGCAACGAGATTGCTACTCGAAGTGAGCCAAT	65	
QY	2132	CCAAGGAGCACCGCAGCTTTTATCACCTTTCTTACAGAGACTCTTTGGGCCCTTTGCTGGAGG	2191	
DB	64	CCAAGGAGCACCGCAGCTTATCACCTTCTCTGACAGAGCTCTCTGGGCCCTCTGCTGGAGG	5	
QY	2192	C 2192		
DB	4	C 4		

RESULT 15

US-09-814-353-16182/c  
; Sequence 16182, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John

```

; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16182
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-16182

Query Match      10.8%; Score 268.6; DB 13; Length 379;
Best Local Similarity 96.8%; Pred. No. 1.1e-74; Indels 1; Gaps 1;
Matches 285; Conservative 0; Mismatches 9;

QY      1438  TGTGCCATTATGTCCACACACATTGTGTGGCTTGCTGCTCTCTACAGACACAGGCAGGGA 1497
Db      379  TGTGCCATTATGTCCACACACATTGTGTGGCTTGCTGCTCTCTACAGACACAGGCAGGGA 320

QY      1498  ATTGATCTCTCCACATTGTCGAAGACTCTTTTGTGATGAAGAGGAAAGTCCTGGGCTCGT 1557
Db      319  AATGATCTCTCCACATTGTCGAAGACTCTTTTGTGATGAAGAGGAAAGTCCTGGGCTAGT 260

QY      1558  GATTTTGACCTGGGGTCTCAGAAATTCAGAAATGCTAGTAATGCATGCCATACAGCTG 1617
Db      259  AATATTGACCTGGGGTCTCAGAAATTCAGAAATGCTAGTAATGCATGCCATACAGCTG 200

QY      1618  CTGGGAAATTGTCTCAAAATCACCCACACTAGCAGGAACGATCAGATTTTTTATCACCCTCC 1677
Db      199  CTGGGAAATTGTCTCAAAATCACCC-CACTAGCAGGAATGATCAGATTTTTTATCACCCTCC 141

QY      1678  AGCACAACTGTGCCATCAGCTTTTCGAACCTCAACTTCTACAGCAATGGGGTACTTTC 1732
Db      140  AGCACAACTGTGCCATCAGCTTTTCGAACCTCAACTTCTACAGCAATGGGGTACTTTC 86

```

Search completed: January 10, 2004, 10:32:05  
Job time : 577.937 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 18:12:45 ; Search time 5946.75 Seconds  
(without alignments)  
17088.240 Million cell updates/sec

Title: US-09-935-290-3  
Perfect score: 2484  
Sequence: 1 atggatgaatctgcactgac.....ttctgagttttgtggtgctg 2484

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.pln.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2484	100.0	2484	6	AX420455 Sequence
2	2484	100.0	3003	6	AX420453 Sequence
3	2482.4	99.9	2487	6	AX431391 Sequence
4	2480.8	99.8	3210	6	AX451419 Sequence
5	2479.2	99.8	3490	9	BC030783 Homo sapi
6	2479.2	99.8	4918	9	HSM804404
7	2477.6	99.7	6387	9	HSM804372
8	2111.8	85.0	5689	9	HSM803772
9	1984.4	79.9	5674	9	AB046780 Homo sapi
10	1879.2	75.7	2646	6	AX401972 Sequence
11	1879.2	75.7	2646	10	AF021348 Rattus no
12	1860	74.9	3857	10	BC019201 Mus muscu
13	1856.8	74.8	6634	10	MUSG3PAT
14	1803	72.6	2519	10	RNU36771
15	267.4	10.8	196657	9	AL391986 Human DNA
16	223.6	9.0	255430	2	AC099387 Rattus no
17	218.2	8.8	197741	2	AC110206 Mus muscu
18	173.8	7.0	206914	2	AC104911 Mus muscu
19	142.6	5.7	4201	4	AF469047 Bos tauru
20	131.6	5.3	147242	2	AC135152 Rattus no
21	123.2	5.0	147242	2	AC135152 Rattus no
22	109	4.4	2683	6	AX710670 Sequence
23	109	4.4	2755	6	AX451420 Sequence
24	107.8	4.3	2382	6	AX710668 Sequence
25	107.8	4.3	2388	6	AX710666 Sequence
26	107.8	4.3	2665	6	AX710664 Sequence
27	103.8	4.2	3450	3	AF145635 Drosophil
28	103.8	4.2	76299	2	AC019585 Drosophil
29	103.8	4.2	172363	3	AC009388 Drosophil
30	103.8	4.2	206190	3	AE003763 Drosophil
31	95.4	3.8	539	11	G98364 S210P6250FD
32	85.6	3.4	38145	3	U64847 Caenorhabdi
33	75	3.0	40532	3	CBRG36D18
34	64.8	2.6	301995	1	AE016779 Pseudomon
35	61.4	2.5	10744	1	AE013659 Yersinia
36	61.4	2.5	214050	1	AJ414142 Yersinia
37	59.2	2.4	300029	1	AE012554 Xylella f
38	58	2.3	310029	1	AE016861 Pseudomon
39	56.6	2.3	299938	1	AE016800 Vibrio vu
40	56	2.3	11392	1	AE003940 Xylella f
41	55	2.2	10802	1	AE015425 Shigella
42	55	2.2	292497	1	AE016989 Shigella
43	54.2	2.2	12847	1	AE012079 Xanthomon
44	53.4	2.1	1141	6	AX083744 Sequence
45	53	2.1	12739	1	AE004787 Pseudomon

ALIGNMENTS

RESULT 1	AX420455	Sequence 3 from Patent WO0216592.	2484 bp	DNA	linear	PAT 18-JUN-2002
AX420455	LOCUS					
DEFINITION	Sequence 3 from Patent WO0216592.					
ACCESSION	AX420455					
VERSION	AX420455.1	GI:21524604				
KEYWORDS						
SOURCE						
ORGANISM	Homo sapiens (human)					
	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE						
AUTHORS	Kapeller-Libermann,R. and Logan,T.J..					
TITLE	Atcr-1, a human acyltransferase and uses thereof					
JOURNAL	Patent: WO 0216592-A 3 28-FEB-2002;					



[illegible]

RESULT	2
LOCUS	AX420453
DEFINITION	Sequence 1 from Patent WO0216592.
ACCESSION	AX420453
VERSION	AX420453.1 GI:21524602
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. . . . Kapeller-Libermann,R. and Logan,T.J. Aacr-1, a human acyltransferase and uses thereof Patent: WO 0216592-A 1 28-FEB-2002; Millennium Pharmaceuticals, Inc. (US)  
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
PATENT NO.	
DNA	linear
SIZE	3003 bp
CHECKSUM	WU0216592
DATE	PAT 18-JUN-2002

BASE COUNT	826 a	683 c	698 g	794 t	2 others
ORIGIN	LLSNGETIISLPQOTFYQVCHETVGKFIQYIGILTVAEHDQDEIDSPSLAEQWMDKLEPE PLSMRDEDESDFEORDCYVKSQSHQOYITFLORLGLPILLEAVYSSAIFVH NFSGPVPEYDEYLOKLHXYLTIRNVAVVAESATYCLVKNAVMPKIDIGVFKETKQK RVLSLSSTFLPQCNKRLLEILVSPVL"				
Query Match	100.0%; Score 2484; DB 6; Length 3003;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 2484; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	ATCGATGAATCTGCACGTGACCCCTTGGTACAATAGATGTTCTTATCTGCGCACATTCATCA	60		
DB	341	ATCGATGAATCTGCACGTGACCCCTTGGTACAATAGATGTTCTTATCTGCGCACATTCATCA	400		
QY	61	GAATACAGTGTTCGTGATGCTAAGCACACAAAGTGAGGAATGGGTGAGTGTGGCTTTAGA	120		
DB	401	GAATACAGTGTTCGTGATGCTAAGCACACAAAGTGAGGAATGGGTGAGTGTGGCTTTAGA	460		
QY	121	CCCACCGTCTTCAGATCTGCAACTTTTAAATGAAAGAAAGCCCTAATGAGTCGGAAGG	180		
DB	461	CCCACCGTCTTCAGATCTGCAACTTTTAAATGAAAGAAAGCCCTAATGAGTCGGAAGG	520		
QY	181	CAATTTGTTGGAAGATGTTGTTACTCTGTCGACCTCCCGAGAGCTGGGACAAATTTTCAAC	240		
DB	521	CCATTTGTTGGAAGATGTTGTTACTCTGTCGACCTCCCGAGAGCTGGGACAAATTTTCAAC	580		
QY	241	CCCAGTATCCGTCCTTTGGTTCGCGAATGTTATTATATCAATGAACTCACACAAGA	300		
DB	581	CCCAGTATCCGTCCTTTGGTTCGCGAATGTTATTATATCAATGAACTCACACAAGA	640		
QY	301	CACCGCGATGGCTTGCAAGACGCCCTTCTTACGTTCTTTTAAATCAAGAGCGAGATGTG	360		
DB	641	CACCGCGATGGCTTGCAAGACGCCCTTCTTACGTTCTTTTAAATCAAGAGCGAGATGTG	700		
QY	361	CATAAGGGCATGTTTGCACCAATGTGACTGAAAAATGTGCTGAACAGCAGTAGAGTCAA	420		
DB	701	CATAAGGGCATGTTTGCACCAATGTGACTGAAAAATGTGCTGAACAGCAGTAGAGTCAA	760		
QY	421	GAGGCAATTTGCAAGTGGCTGCTGAATTAACCCCTGATGGTTCGCCACGACGCAATCA	480		
DB	761	GAGGCAATTTGCAAGTGGCTGCTGAATTAACCCCTGATGGTTCGCCACGACGCAATCA	820		
QY	481	AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTTAAAGGATCTTCAAGAAATGGTTCACCT	540		
DB	821	AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTTAAAGGATCTTCAAGAAATGGTTCACCT	880		
QY	541	GTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTTAAACTGTTCAAACAGCTTC	600		
DB	881	GTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTTAAACTGTTCAAACAGCTTC	940		
QY	601	TTTTGGAAACATTCAAAATTCACAAAGTGCACCTTGAATGGTTAAAGCTGCAACTGAGACG	660		
DB	941	TTTTGGAAACATTCAAAATTCACAAAGTGCACCTTGAATGGTTAAAGCTGCAACTGAGACG	1000		
QY	661	AATTTCCCGCTTCGTGTTCTACAGTTCATAGATCCCATATTCGACTATCTGCTGCTCACT	720		
DB	1001	AATTTCCCGCTTCGTGTTCTACAGTTCATAGATCCCATATTCGACTATCTGCTGCTCACT	1060		
QY	721	TTCAATTCCTTCTGCGCAATCAACATCAAAGCACCATACATTCGCTTCAGGCAATAATCTCAAC	780		
DB	1061	TTCAATTCCTTCTGCGCAATCAACATCAAAGCACCATACATTCGCTTCAGGCAATAATCTCAAC	1120		
QY	781	ATCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTCATACAGCAAGGCTC	840		
DB	1121	ATCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTCATACAGCAAGGCTC	1180		
QY	841	GATGAAACACAGATGGAACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA	900		
DB	1181	GATGAAACACAGATGGAACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA	1240		
QY	901	GTTGAATTAATTCGACAGCAGCAATTTCTTGGAGATCTTCTCTGGAAGGCACACGTTCTTAGG	960		

Db 1241 GTTGAATTAATCTCGACAGCAATCTTGGAGATCTTCTGGAAGCCACACGCTTCAGG 1300  
QY 961 AGTGGAAAAACCTCTTGTGCTCGGGCAGGACTTTTGTGTCAGTGTGGTAGATACTCTGTCT 1020  
Db 1301 AGTGGAAAAACCTCTTGTGCTCGGGCAGGACTTTTGTGTCAGTGTGGTAGATACTCTGTCT 1360  
QY 1021 ACCAATGTCATCCAGACATCTTTGATAATACCTGTGTGGAATCTCTATGATCGCATTC 1080  
Db 1361 ACCAATGTCATCCAGACATCTTTGATAATACCTGTGTGGAATCTCTATGATCGCATTC 1420  
QY 1081 GAAGGTCACTACAAATGTAAGTGAACCTGGGCAACCTTAAGAAATGAGAGCCCTGTGAGT 1140  
Db 1421 GAAGGTCACTACAAATGTAAGTGAACCTGGGCAACCTTAAGAAATGAGAGCCCTGTGAGT 1480  
QY 1141 GTAGCAAGAGGTGTTATTGAATGTTACGAAAAAATATGTTGTGTGTCGAGTGGATTTT 1200  
Db 1481 GTAGCAAGAGGTGTTATTGAATGTTACGAAAAAATATGTTGTGTGTCGAGTGGATTTT 1540  
QY 1201 GCACAGCCATTTTCTTAAAGGAATATTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCT 1260  
Db 1541 GCACAGCCATTTTCTTAAAGGAATATTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCT 1600  
QY 1261 CTACTTTCCCTGGAGCAAGCTTGTACAGCTATATCTTCCCTTCAGAACCCAGTGTGCT 1320  
Db 1601 CTACTTTCCCTGGAGCAAGCTTGTACAGCTATATCTTCCCTTCAGAACCCAGTGTGCT 1660  
QY 1321 GCTGATGAAGGTAGAGACAGTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db 1661 GCTGATGAAGGTAGAGACAGTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1720  
QY 1381 CGAAGGAGGTGATGTCGAAATCTGGCTGAGCATATCTTATCTGCTAGCAAGTCTGT 1440  
Db 1721 CGAAGGAGGTGATGTCGAAATCTGGCTGAGCATATCTTATCTGCTAGCAAGTCTGT 1780  
QY 1441 GCATATATGTCACACATGTTGGCTTGTGCTGCTCTCTACAGACACAGGCGAGGAAT 1500  
Db 1781 GCATATATGTCACACATGTTGGCTTGTGCTGCTCTCTACAGACACAGGCGAGGAAT 1840  
QY 1501 GATCTCTCCACATTTGGTGAAGACTTCTTTGTGATGAAGAGGAAGTCTGGCTCGTGAT 1560  
Db 1841 GATCTCTCCACATTTGGTGAAGACTTCTTTGTGATGAAGAGGAAGTCTGGCTCGTGAT 1900  
QY 1561 TTTGACCTGGGGTCTCAGAAATTCAGAAATGTAGTAATGATGTCATACAGCTGCTG 1620  
Db 1901 TTTGACCTGGGGTCTCAGAAATTCAGAAATGTAGTAATGATGTCATACAGCTGCTG 1960  
QY 1621 GGAATTTGTGTCACATCACCACACTAGCAGGAAAGATGAGTTTTTTATCACCACCCAGC 1680  
Db 1961 GGAATTTGTGTCACATCACCACACTAGCAGGAAAGATGAGTTTTTTATCACCACCCAGC 2020  
QY 1681 ACAACTGTCCCATCAGTCTTCGAACTCAACTTCTACAGCAATGGGTACTTTCATGCTTT 1740  
Db 2021 ACAACTGTCCCATCAGTCTTCGAACTCAACTTCTACAGCAATGGGTACTTTCATGCTTT 2080  
QY 1741 ATCATGAGGCCATCATAGCTTTCAGCTTTTATGCACTTCTGAACAGAGGGGACTGGGG 1800  
Db 2081 ATCATGAGGCCATCATAGCTTTCAGCTTTTATGCACTTCTGAACAGAGGGGACTGGGG 2140  
QY 1801 GGTCCCATAGACCCACCTAATCTGATCAGCAGAGCAGCTGTGGGAAAGGGCGCC 1860  
Db 2141 GGTCCCATAGACCCACCTAATCTGATCAGCAGAGCAGCTGTGGGAAAGGGCGCC 2200  
QY 1861 AGCTGTGCTACCTTCTCCATGAGGCAACCTCTCAGCTGCTGCCAGACATTTTAC 1920  
Db 2201 AGCTGTGCTACCTTCTCCATGAGGCAACCTCTCAGCTGCTGCCAGACATTTTAC 2260  
QY 1921 CAAAGTCTGCCATGAACAGTAGGAAGTTTATCCAGTATGGCATTTTACAGTGGCAGAG 1980  
Db 2261 CAAAGTCTGCCATGAACAGTAGGAAGTTTATCCAGTATGGCATTTTACAGTGGCAGAG 2320  
QY 1981 CACGATGACGAGGAATATCAGTCTGCTGCTGAGCAGCAGTGGGCAAGAGCTTT 2040  
Db 2321 CACGATGACGAGGAATATCAGTCTGCTGCTGAGCAGCAGTGGGCAAGAGCTTT 2380

QY 2041 CCAGAACCTTTGTCTTGGAGAAAGTGAAGAGATGAAGACAGTGAATTTGGGGAGGAA 2100  
Db 2381 CCAGAACCTTTGTCTTGGAGAAAGTGAAGAGATGAAGACAGTGAATTTGGGGAGGAA 2440  
QY 2101 CAGCGAGATGCTTACTGAAGGTGAGCCCAATCAAGGAGCAGCAGAGTTTATCAGCTTC 2160  
Db 2441 CAGCGAGATGCTTACTGAAGGTGAGCCCAATCAAGGAGCAGCAGAGTTTATCAGCTTC 2500  
QY 2161 TTACAGAGACTCCTTGGGCTTTTGTGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2220  
Db 2501 TTACAGAGACTCCTTGGGCTTTTGTGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2560  
QY 2221 AACTTCAGTGGTCTCTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTAAATA 2280  
Db 2561 AACTTCAGTGGTCTCTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTAAATA 2620  
QY 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATTTGTTGTAAG 2340  
Db 2621 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATTTGTTGTAAG 2680  
QY 2341 AATGCTGTGAAATGTTTAAAGGATATTTGGGGTTTCAAGGAGACCAACAAAGAGAGTG 2400  
Db 2681 AATGCTGTGAAATGTTTAAAGGATATTTGGGGTTTCAAGGAGACCAACAAAGAGAGTG 2740  
QY 2401 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAAACCTTTCTAGAA 2460  
Db 2741 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAAACCTTTCTAGAA 2800  
QY 2461 TATATTTCTGAGTTTTGTTGTTGCTG 2484  
Db 2801 TATATTTCTGAGTTTTGTTGTTGCTG 2824

## RESULT 3

AX431391  
LOCUS AX431391 2487 bp DNA linear PAT 28-JUN-2002  
DEFINITION Sequence 24 from Patent WO0240666.  
ACCESSION AX431391  
VERSION AX431391.1 GI:21656238  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Enkins,D.K., Winther,M.D., Haardt,M., Goldberg,Y.P., Nwaka,S.O.,  
Ponton,A., Allen,S.J., de Antueno,R.J. and Knickle,L.C.  
TITLE Fat regulated genes, uses thereof, and compounds for modulating  
same  
JOURNAL Patent: WO 0240666-A 24 23-MAY-2002;  
XENON GENETICS INC (CA)  
FEATURES  
Location/Qualifiers  
source 1..2487  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 706 a 548 c 570 g 663 t

Query Match 99.9%; Score 2482.4; DB 6; Length 2487;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2483; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGATGAATCTGCACTGACCTTGGTACAATAGATGTTTCTTATCTGCCACATTCATCA 60

Db 1 ATGGATGAATCTGCACTGACCTTGGTACAATAGATGTTTCTTATCTGCCACATTCATCA 60

QY 61 GAATACAGTGTGTCGATGTAAGCACACAGTGGGAATGGGTGAGTGGCTTTTGA 120

Db 61 GAATACAGTGTGTCGATGTAAGCACACAGTGGGAATGGGTGAGTGGCTTTTGA 120

QY 121 CCCACGCTTTCAGATCTCAACTTTTAAATGGAAGAAAGCCTAATAGTTCGGAAGG 180

Db 121 CCAACCATCTTCAGATCTGCAACTTTAAATGAAAGAAAGCCTAATGAGTCGGAAAGG 180  
Qy 181 CCAATTTGTTGAAGATGTTGTTACTCTGCACTCCCCAGAGCTGGGACAAAATTTTCAAC 240  
Db 181 CCAATTTGTTGAAGATGTTGTTACTCTGCACTCCCCAGAGCTGGGACAAAATTTTCAAC 240  
Qy 241 CCAGTATCCCGTCTTTGGGTTTGGGAAATGTTATTTATATCAATGAAACTCACACAAGA 300  
Db 241 CCAGTATCCCGTCTTTGGGTTTGGGAAATGTTATTTATATCAATGAAACTCACACAAGA 300  
Qy 301 CACCCGGATGCTTTGCAAGACGCTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTG 360  
Db 301 CACCCGGATGCTTTGCAAGACGCTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTG 360  
Qy 361 CATAAGGGCATGTTTGGCCACCAATGTGACTGAAAAATGTCTGAAACAGCAGTAGAGTACAA 420  
Db 361 CATAAGGGCATGTTTGGCCACCAATGTGACTGAAAAATGTCTGAAACAGCAGTAGAGTACAA 420  
Qy 421 GAGGCAATTCGAAAGTGGTCTGCTGAAATTAACCCCTGATGTTCTGCCAGCAGCAATCA 480  
Db 421 GAGGCAATTCGAAAGTGGTCTGCTGAAATTAACCCCTGATGTTCTGCCAGCAGCAATCA 480  
Qy 481 AAAGCCGTTTAAACAAGTGAAGAAAGAAAGCTTAAAGGATTTCTCAAGAAATGGTTGCCACT 540  
Db 481 AAAGCCGTTTAAACAAGTGAAGAAAGAAAGCTTAAAGGATTTCTCAAGAAATGGTTGCCACT 540  
Qy 541 GTCTCACCGGCAATGATCAGACTGACTGGGTGGTGGTCTGCTTAAAGGATTTCTCAAGAAATGGTTGCCACT 600  
Db 541 GTCTCACCGGCAATGATCAGACTGACTGGGTGGTGGTCTGCTTAAAGGATTTCTCAAGAAATGGTTGCCACT 600  
Qy 601 TTTTGGAACTTCAAAATTCACAAAGTCAAACTTGAGATGGTTTAAAGCTGCAACTGAGACG 660  
Db 601 TTTTGGAACTTCAAAATTCACAAAGTCAAACTTGAGATGGTTTAAAGCTGCAACTGAGACG 660  
Qy 661 AATTTGCCGCTTCTGTTTCTACCACTCATAGATCCCATATTTGACTATCTGCTGCTCACT 720  
Db 661 AATTTGCCGCTTCTGTTTCTACCACTCATAGATCCCATATTTGACTATCTGCTGCTCACT 720  
Qy 721 TTCACTTCTTCTGTCATAAATCAAAAGCACCACATATGTTTCAAGGCAATAATCTCAAC 780  
Db 721 TTCACTTCTTCTGTCATAAATCAAAAGCACCACATATGTTTCAAGGCAATAATCTCAAC 780  
Qy 781 ATCCCAATCTTCAGTACCTTGATCATAAGCTTTGGGGCTTCTTCATACGACGAGGCTC 840  
Db 781 ATCCCAATCTTCAGTACCTTGATCATAAGCTTTGGGGCTTCTTCATACGACGAGGCTC 840  
Qy 841 GATGAAACACAGATGGAGGAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db 841 GATGAAACACAGATGGAGGAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Qy 901 GTTGAAATTTACTTCGACAGCAGCAATTTCTTGAGATCTTCTGGAAGGCACACGTTCTAGG 960  
Db 901 GTTGAAATTTACTTCGACAGCAGCAATTTCTTGAGATCTTCTGGAAGGCACACGTTCTAGG 960  
Qy 961 AGTGGAAAAACCTCTTGTGCTCGGCAGGACTTTTGTGAGTGTGGTAGACTCTGTCT 1020  
Db 961 AGTGGAAAAACCTCTTGTGCTCGGCAGGACTTTTGTGAGTGTGGTAGACTCTGTCT 1020  
Qy 1021 ACCAATGTCATCCAGACATCTTGATATAATCTGTTGGAATCTCTATATGCGCATATC 1080  
Db 1021 ACCAATGTCATCCAGACATCTTGATATAATCTGTTGGAATCTCTATATGCGCATATC 1080  
Qy 1081 GAAGTCTACTACAATGGTGAACAATCTGGGCAAACTTAAGAGAAATGAGAGCTGTGGAGT 1140  
Db 1081 GAAGTCTACTACAATGGTGAACAATCTGGGCAAACTTAAGAGAAATGAGAGCTGTGGAGT 1140  
Qy 1141 GTAGCAAGAGGTGTTATTAGATGTTACGAAAAAACTATGGTTGTGTCGAGTGGATTTT 1200  
Db 1141 GTAGCAAGAGGTGTTATTAGATGTTACGAAAAAACTATGGTTGTGTCGAGTGGATTTT 1200  
Qy 1201 GCACAGCCATTTCTTAAAGGAATTTTAAAGGCCAAAGTCAAGAACCCGGTGTCTGCT 1260

Db 1201 GCACAGCCATTTTCTTAAAGGAATATTTTAAAGGCCAAAGTCAAGAACCCGGTGTCTGCT 1260  
Qy 1261 CTACTTTCCCTGGAGCAAGCGTGTGTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320  
Db 1261 CTACTTTCCCTGGAGCAAGCGTGTGTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320  
Qy 1321 GCTGATGAAGGTGAGACACGTCCTCAATTAATAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db 1321 GCTGATGAAGGTGAGACACGTCCTCAATTAATAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Qy 1381 CGAAGGAGGTGATTTGCAAAATCTGGCTGAGCATATCTTATTTCACTGTAGCAAGTCTGT 1440  
Db 1381 CGAAGGAGGTGATTTGCAAAATCTGGCTGAGCATATCTTATTTCACTGTAGCAAGTCTGT 1440  
Qy 1441 GCCAATATGTCCACACACATTTGTGCTTGTGCTCTCTACAGACACAGGAGGAATTT 1500  
Db 1441 GCCAATATGTCCACACACATTTGTGCTTGTGCTCTCTCTACAGACACAGGAGGAATTT 1500  
Qy 1501 GATCTCTCCATTTGGTTCGAAGACTTCTTTGTGATGAAAGGAAGTCTTGGCTCGTGT 1560  
Db 1501 GATCTCTCCATTTGGTTCGAAGACTTCTTTGTGATGAAAGGAAGTCTTGGCTCGTGT 1560  
Qy 1561 TTTGACCTGGGGTCTTCAGGAAATTCAGAAATTCAGAAATTCAGTAAATGTCATGCTGCTG 1620  
Db 1561 TTTGACCTGGGGTCTTCAGGAAATTCAGAAATTCAGAAATTCAGTAAATGTCATGCTGCTG 1620  
Qy 1621 GGAATTTGTGTCACAATCACCACACTAGCAGGAACGATGAGTTTTTTATCACCCCCAGC 1680  
Db 1621 GGAATTTGTGTCACAATCACCACACTAGCAGGAACGATGAGTTTTTTATCACCCCCAGC 1680  
Qy 1681 ACAACTGTCCCATCAGTCTTCGAACCTTCTACAGCAATGGGGTACTTCTGTCTTTT 1740  
Db 1681 ACAACTGTCCCATCAGTCTTCGAACCTTCTACAGCAATGGGGTACTTCTGTCTTTT 1740  
Qy 1741 ATCATGAGGCGCATCATAGCTTTGAGCCTTTATGCGAGTTCTGAAACAGAGGGGACTGGGG 1800  
Db 1741 ATCATGAGGCGCATCATAGCTTTGAGCCTTTATGCGAGTTCTGAAACAGAGGGGACTGGGG 1800  
Qy 1801 GGTCCCACTAGCACCCCACTTAACCTCATCAGCAGGAGCAGCTGGTCCGGAAGCGGCC 1860  
Db 1801 GGTCCCACTAGCACCCCACTTAACCTCATCAGCAGGAGCAGCTGGTCCGGAAGCGGCC 1860  
Qy 1861 AGCTGTGTACCTTCTCTCCAATGAAGGCAACCATCTCACTGCTTCCGACAGACATTTTAC 1920  
Db 1861 AGCTGTGTGTACCTTCTCTCCAATGAAGGCAACCATCTCACTGCTTCCGACAGACATTTTAC 1920  
Qy 1921 CAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTCTTACAGTGGCAGAG 1980  
Db 1921 CAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTCTTACAGTGGCAGAG 1980  
Qy 1981 CACGATCACAGGAAGATATCAGTCTTAGTCTTGTGCTGAGCAGCAGTGGGACAGAGCTT 2040  
Db 1981 CACGATCACAGGAAGATATCAGTCTTAGTCTTGTGCTGAGCAGCAGTGGGACAGAGCTT 2040  
Qy 2041 CCAGAACTTTGTCTTGGAGAAAGTGAAGAAAGATGAAGACAGTGAATTTTGGGGAGGAA 2100  
Db 2041 CCAGAACTTTGTCTTGGAGAAAGTGAAGAAAGATGAAGACAGTGAATTTTGGGGAGGAA 2100  
Qy 2101 CAGGAGATTTGCTACTTGAAGGTGAGCCAAATCCAAAGGAGCACAGAGTTTATCACCTTC 2160  
Db 2101 CAGGAGATTTGCTACTTGAAGGTGAGCCAAATCCAAAGGAGCACAGAGTTTATCACCTTC 2160  
Qy 2161 TTACAGAGACTCTTGGGGCTTGTGAGGCGCTTACAGCTCTGCTGCAATCTTTGTTTAC 2220  
Db 2161 TTACAGAGACTCTTGGGGCTTGTGAGGCGCTTACAGCTCTGCTGCAATCTTTGTTTAC 2220  
Qy 2221 AACTTCAAGTGTCTGTTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAATA 2280  
Db 2221 AACTTCAAGTGTCTGTTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAATA 2280  
Qy 2281 ACCAGAACAGAAAGAAATGTTGCGAGTATATGCTGAGAGTGCCACATATGTTCTTGTGAAG 2340  
Db 2281 ACCAGAACAGAAAGAAATGTTGCGAGTATATGCTGAGAGTGCCACATATGTTCTTGTGAAG 2340



QY 2341 AATGCTGTGAAATGTTTAAAGGATATTGGGTTTCAAGGAGACCAAAACAAAAGAGAGTG 2400  
Db |||||  
2341 AATGCTGTGAAATGTTTAAAGGATATTGGGTTTCAAGGAGACCAAAACAAAAGAGAGTG 2400  
QY 2401 TCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAATGCAACCGACAAACAACTTTAGAA 2460  
Db |||||  
2401 TCTGTTTTAGAACTGAGCAGCAGCTTTTCTACCTCAATGCAACCGACAAACAACTTTAGAA 2460  
QY 2461 TATATTCTGAGTTTGTGGTGCTG 2484  
Db |||||  
2461 TATATTCTGAGTTTGTGGTGCTG 2484

RESULT 4  
AX451419  
LOCUS AX451419 3210 bp DNA linear PAT 03-JUL-2002  
DEFINITION Sequence 26 from Patent WO0226950.  
ACCESSION AX451419  
VERSION AX451419.1 GI:21698428  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Lal,P.G., Tang,Y.T., Yue,H., Burford,N., Gandhi,A.R., Warren,B.A.,  
Yao,M.G., Tribouley,C.M., Baughn,M.R., Lee,E.A., Hafalia,A.J.,  
Lu,Y., Griffin,J.A., Sanjanwala,M.S. and Ding,L.  
Transferrases  
TITLE Patent: WO 0226950-A 26 04-APR-2002;  
JOURNAL Incyte Genomics, Inc. (US)  
FEATURES  
source location/Qualifiers  
1..3210  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/notes="Incyte ID No: 1517230CB1"  
BASE COUNT 896 a 721 c 735 g 858 t  
ORIGIN

Query Match 99.9%; Score 2480.8; DB 6; Length 3210;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2482; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGATGAATCTGCATGACCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 60  
Db |||||  
338 ATGGATGAATCTGCATGACCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 397  
QY 61 GAATACAGTCTTGGTTCGATGAAGCACACAAGTGAGGATGGGTTGAGTGTGGCTTTAGA 120  
Db |||||  
398 GAATACAGTGTGGTTCGATGAAGCACACAAGTGAGGATGGGTTGAGTGTGGCTTTAGA 457  
QY 121 CCCACCGCTTTCAGATCTGCAACTTTAAATGGAAAGAAAGCCCTAATGAGTCGGAAGG 180  
Db |||||  
458 CCCACCATCTTCAGATCTGCAACTTTAAATGGAAAGAAAGCCCTAATGAGTCGGAAGG 517  
QY 181 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGCTGGGACAAATTTTCAAC 240  
Db |||||  
518 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGCTGGGACAAATTTTCAAC 577  
QY 241 CCCAGTATCCCGTCTTGGGTTGGCGAATGTTTATATCAATGAAACTCACACAAGA 300  
Db |||||  
578 CCCAGTATCCCGTCTTGGGTTGGCGAATGTTTATATCAATGAAACTCACACAAGA 637  
QY 301 CACCGGGATGGCTTGCAGACGCTTTCTTACGTTCTTTTATTCAGAGCGAGATGTG 360  
Db |||||  
638 CACCGGGATGGCTTGCAGACGCTTTCTTACGTTCTTTTATTCAGAGCGAGATGTG 697  
QY 361 CATAAGGCGATGTTTGGCACCATGTGACGTGAAATGTGCTGAACAGCAGTAGAGTACAA 420  
Db |||||  
698 CATAAGGCGATGTTTGGCACCATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 757

QY 421 GAGGCAATTGCAGAGTGGCTGCTGAAATTAACCCCTGATGGTTCTGCCAGAGCAATCA 480  
Db |||||  
758 GAGGCAATTGCAGAGTGGCTGCTGAAATTAACCCCTGATGGTTCTGCCAGAGCAATCA 817  
QY 481 AAAGCGTTTAAACAAAGTGAAGAAAGCTAAAGGATTTCTTCAAGAAATGTTGCCACT 540  
Db |||||  
818 AAAGCGTTTAAACAAAGTGAAGAAAGCTAAAGGATTTCTTCAAGAAATGTTGCCACT 877  
QY 541 GTCTCACCGGCAATGATCAGACTGACTGGTGGGTGCTGCTAAACCTGTTTCAACAGCTTC 600  
Db |||||  
878 GTCTCACCGGCAATGATCAGACTGACTGGTGGGTGCTGCTAAACCTGTTTCAACAGCTTC 937  
QY 601 TTTTGGAACTTCAAAATTCACAAAGGTCAAATTGAGATGGTTAAAGCTCCAACTGAGACG 660  
Db |||||  
938 TTTTGGAACTTCAAAATTCACAAAGGTCAAATTGAGATGGTTAAAGCTCCAACTGAGACG 997  
QY 661 AATTTGCGCGTCTGTTTCTACAGTTTATAGATCCCATATTTGACTATCTGCTCAGT 720  
Db |||||  
998 AATTTGCGCGTCTGTTTCTACAGTTTATAGATCCCATATTTGACTATCTGCTCAGT 1057  
QY 721 TTTCAATTTCTTCTGCCATAAATCAAGCACCACATCATTTGCTTCAAGGCAATATCTCAAC 780  
Db |||||  
1058 TTTCAATTTCTTCTGCCATAAATCAAGCACCACATCATTTGCTTCAAGGCAATATCTCAAC 1117  
QY 781 ATCCCAATCTTCAGTACCTTGTATCCATAGCTTGGGGGCTTCTTATACAGCAAGGCTC 840  
Db |||||  
1118 ATCCCAATCTTCAGTACCTTGTATCCATAGCTTGGGGGCTTCTTATACAGCAAGGCTC 1177  
QY 841 GATGAAACACAGATGGAGGAAAGATGTTCTATAGAGCTTTGCTTCCATGGGCATATA 900  
Db |||||  
1178 GATGAAACACAGATGGAGGAAAGATGTTCTATAGAGCTTTGCTTCCATGGGCATATA 1237  
QY 901 GTTGAATTTACTTCGACAGCAGCAATTTCTTGGAGATCTTCTCGAAGGACACGTTCTAGG 960  
Db |||||  
1238 GTTGAATTTACTTCGACAGCAGCAATTTCTTGGAGATCTTCTCGAAGGACACGTTCTAGG 1297  
QY 961 AGTGGAAAAACCTTCTGTGCTCGGCGAGGACTTTTGTTCAGTTGTGGTAGATACTCTGTCT 1020  
Db |||||  
1298 AGTGGAAAAACCTTCTGTGCTCGGCGAGGACTTTTGTTCAGTTGTGGTAGATACTCTGTCT 1357  
QY 1021 ACCAATGTCATCCAGACATCTTGTATATACCTGTTGGAATCTCTATGATCGCATATC 1080  
Db |||||  
1358 ACCAATGTCATCCAGACATCTTGTATATACCTGTTGGAATCTCTATGATCGCATATC 1417  
QY 1081 GAAGGTCACTCAATGTTGAACAACTGGGCAAACTTAAGAAAGTGAAGCTGTGGAGT 1140  
Db |||||  
1418 GAAGGTCACTCAATGTTGAACAACTGGGCAAACTTAAGAAAGTGAAGCTGTGGAGT 1477  
QY 1141 GTAGCAAGAGTGTATTAGAAATGTTACGAAAAAACTATGTTGTCTCCGAGTGGATTTT 1200  
Db |||||  
1478 GTAGCAAGAGTGTATTAGAAATGTTACGAAAAAACTATGTTGTCTCCGAGTGGATTTT 1537  
QY 1201 GCACAGCCATTTTCTTAAAGGAATATTTAGAAAGCAAAAGTCAGAAACCGGTGTCTGCT 1260  
Db |||||  
1538 GCACAGCCATTTTCTTAAAGGAATATTTAGAAAGCAAAAGTCAGAAACCGGTGTCTGCT 1597  
QY 1261 CTACTTTCCCTCGGACAGCGTTGTTACAGCTATCTTCTTCAAGCCCCAGTGTGCT 1320  
Db |||||  
1598 CTACTTTCCCTCGGACAGCGTTGTTACAGCTATCTTCTTCAAGCCCCAGTGTGCT 1657  
QY 1321 GCTGATGAAGGTAGAGACACGTCCTCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db |||||  
1658 GCTGATGAAGGTAGAGACACGTCCTCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1717  
QY 1381 CGAAGAGGTTGAATGCAAAATCTGGCTGAGCATATTTCTATTCTGCTAGCAAGTCTGCT 1440  
Db |||||  
1718 CGAAGAGGTTGAATGCAAAATCTGGCTGAGCATATTTCTATTCTGCTAGCAAGTCTGCT 1777  
QY 1441 GCGATTATGTCACACATTTGCTGCTGCTTCTCTACAGACACAGGCGAGGAAAT 1500  
Db |||||  
1778 GCGATTATGTCACACATTTGCTGCTGCTTCTCTACAGACACAGGCGAGGAAAT 1837  
QY 1501 GATCTCTCCATTTGGTCCGAGACTTTTGTGTGATGAAAGAGGAAAGTCTCTGGCTCGTGAT 1560



179 ATGGATGAATATGCACTGACCCCTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCA 238  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

1141 GTAGCAAGAGGTGTTATTAGAAATGTTAGCAAAAAATATGTTGTTGTCGAGTGGATTTT 1200  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000

QY 2221 AACTTCAGTGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTGCACAAATACCTAATA 2280  
Db |||||||  
QY 2399 AACTTCAGTGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTGCACAAATACCTAATA 2458  
Db |||||||  
QY 2281 ACCAGAACAGAAAGAAAGTTGACGATATCTGAGAGTGCACATATGCTTGTGAG 2340  
Db |||||||  
QY 2459 ACCAGAACAGAAAGAAAGTTGACGATATCTGAGAGTGCACATATGCTTGTGAG 2518  
Db |||||||  
QY 2341 AATGCTGTGAAATGTTTAAAGGATATGGGGTTCCTCAAGGAGACCAAAACAAAGAGAGTG 2400  
Db |||||||  
QY 2519 AATGCTGTGAAATGTTTAAAGGATATGGGGTTCCTCAAGGAGACCAAAACAAAGAGAGTG 2578  
Db |||||||  
QY 2401 TCTGTTTAAAGTGTGAGAGAGTCTTCTACCTCAATGCAACCGACAAACTTCTAGAA 2460  
Db TCTGTTTAAAGTGTGAGAGAGTCTTCTACCTCAATGCAACCGACAAACTTCTAGAA 2638  
QY 2461 TATATTCCTGAGTTTGTGCTGCTG 2484  
Db TATATTCCTGAGTTTGTGCTGCTG 2662

RESULT 6  
HSM804404 HSM804404 4918 bp mRNA linear PRI 30-APR-2003  
LOCUS Homo sapiens mRNA; cDNA DKFZp451P0819 (from clone DKFZp451P0819);  
DEFINITION complete cds.  
ACCESSION AL833093  
VERSION AL833093.2 GI:30268382  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 4918)  
AUTHORS Bloeker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B., Amid,C.,  
Fobo,G., Han,M., Osanger,A. and Wiemann,S.  
TITLE Direct Submission  
JOURNAL Submitted (30-APR-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANY  
COMMENT On Apr 30, 2003 this sequence version replaced gi:21733684.  
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
This clone (DKFZp451P0819) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
information about the clone and the sequencing project is available  
at http://mips.gsf.de/proj/cDNA/  
Location/Qualifiers  
1. 4918  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="10q25.2"  
/clone="DKFZp451P0819"  
/tissue\_type="cDNA-collection"  
/clone\_lib="451 (synonym: hicci). Vector pSport1; host  
DH10B; sites NotI + SalI"  
/dev\_stage="adult"  
1. 4918  
/gene="DKFZp451P0819"  
200..2686  
/gene="DKFZp451P0819"  
/note="glycerol-3-phosphate acyltransferase"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAD8932.1"  
/db\_xref="GI:30268383"  
/translation="MDESALTGLGIDVSLPHSSEYSGVRCKHTSEBWGECGFRTPIF  
RSATLKWKESLMSRKPRGRCYSCPTQSWDKFNPSPISLGRNVIYINETHRHR  
GWLARLLSYLFIQERDVHKGWATNTVENLNSRVQEAIAEVAALNPDGSAQQQS

FEATURES  
source

KAVNKVKAKRILQEMVATVSPAMIRLTGWLKLFNSFWNFQIHKGQLEMVKAAT  
ETNPLLFVHRSHIDYLLTFLFCHNIKAPYIASGNLNIPIFSTLHLKGGFFI  
RRLEDPDKRKDYLRAHLGHVLELROQFLEIPIEGTRSSGKTSACRAGLSV  
VVDTLSTNVPDILIIIPVGISYDRIIGHYNGEOLGPKKNESIWSVARGVIRMRKN  
YGCVRDPAOPFSLKEYLESQSKPVSALLSLEQALLPAILPSRPSDAADRGRTLSIN  
ESNATDESRRRLIANLAHILPTASCSAINTSHIVACLLIHRHQGLDLSLTVED  
FFVMKEVLARDFDLFGSNDSDVMHAIQLLNCVITITHTSRNDEFITPSTVPSV  
FELNFGVGLVHPIMEAIILACSILYAVLNKRGSGPTSTPNLISQBLVRKAASLCY  
LLSNEGTSILPCCOTFYOVCHETVKFIRYGLITVAEHDDOEDI SPSLAEOQWDKLLPE  
PLSWRSEDEEDSDPGEORQDYLKVSQSEHQOQITFLORLLGPLLEAYSSAIFVH  
NFGSPVEPYLQKLKHLITRERNVAVYAESATYCLVKNVAMFKDIDGVFKETQK  
RVSVLELSSFTLPCQNKRLLEYILSFVLL"  
4893  
polyA\_site  
BASE COUNT 1379 a 1005 c 1075 g 1459 t  
ORIGIN  
Query Match 99.8%; Score 2479.2; DB 9; Length 4918;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2481; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 ATGGATGAATCTGCACGTGACCCCTTGGTACATAGATGTTCTTATCTGCCACATTCATCA 60  
Db |||||||  
QY 200 ATGGATGAATCTGCACGTGACCCCTTGGTACATAGATGTTCTTATCTGCCACATTCATCA 259  
Db |||||||  
QY 61 GAATACAGAGTGTGTCGATGTAAGCACACAAGTCAGCAATGGGGTGAGTGGCTTTAGA 120  
Db |||||||  
QY 260 GAATACAGAGTGTGTCGATGTAAGCACACAAGTCAGCAATGGGGTGAGTGGCTTTAGA 319  
Db |||||||  
QY 121 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAAGAAAGCCCTTAATGAGTCGGAAGG 180  
Db |||||||  
QY 320 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAAGAAAGCCCTTAATGAGTCGGAAGG 379  
Db |||||||  
QY 181 CCATTGCTTGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGCAAAATTTTCAAC 240  
Db |||||||  
QY 380 CCATTGCTTGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGCAAAATTTTCAAC 439  
Db |||||||  
QY 241 CCCAGTATCCCGTCTTTGGGGTTTCGGAATGTTTATTTATATCAATGAAATCAACACAAGA 300  
Db |||||||  
QY 440 CCCAGTATCCCGTCTTTGGGGTTTCGGAATGTTTATTTATATCAATGAAATCAACACAAGA 499  
Db |||||||  
QY 301 CACCGCGAGTGGCTTGAAGACGCTTTTCTTACCGTTCTTTTATTTCAAGAGCGAGATGTG 360  
Db |||||||  
QY 500 CACCGCGAGTGGCTTGAAGACGCTTTTCTTACCGTTCTTTTATTTCAAGAGCGAGATGTG 559  
Db |||||||  
QY 361 CATGAAGGCGATGTTTGGCCCAATGTCACGAAATGTCGTAACAGCAGTAGAGTACAA 420  
Db |||||||  
QY 560 CATGAAGGCGATGTTTGGCCCAATGTCACGAAATGTCGTAACAGCAGTAGAGTACAA 619  
Db |||||||  
QY 421 GAGGCAATTGCAAGAGTGGCTGCTGAATTAACCCCTGATGTTCTGCCACGACCAATCA 480  
Db |||||||  
QY 620 GAGGCAATTGCAAGAGTGGCTGCTGAATTAACCCCTGATGTTCTGCCACGACCAATCA 679  
Db |||||||  
QY 481 AAAGCCGTTAAACAAAGTGAAGAAAGAAAGTAAAGGATTTCTTCAAGAAATGGTTGCCACT 540  
Db |||||||  
QY 680 AAAGCCGTTAAACAAAGTGAAGAAAGTAAAGGATTTCTTCAAGAAATGGTTGCCACT 739  
Db |||||||  
QY 541 GTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db |||||||  
QY 740 GTCTCACCGCAATGATCAGACTGACTGGGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 799  
Db |||||||  
QY 601 TTTTGGAAACATTCAAATTCACAAAGTCAACTTCAGATGTTTAAAGTGGTAAAGTGGTAAAGT 660  
Db |||||||  
QY 800 TTTTGGAAACATTCAAATTCACAAAGTCAACTTCAGATGTTTAAAGTGGTAAAGTGGTAAAGT 859  
Db |||||||  
QY 661 AATTGCGCGTCTCTGTTTCTTACAGATTCATAGATCCCATATTTGACTATCTGCTGCTCACT 720  
Db |||||||  
QY 860 AATTGCGCGTCTCTGTTTCTTACAGATTCATAGATCCCATATTTGACTATCTGCTGCTCACT 919  
Db |||||||  
QY 721 TTCAATCTCTTCTGCCCATTAACATCAAAAGCACCACATATTTGCTTTCAGGCAATTAATCTCAAC 780  
Db |||||||  
QY 920 TTCAATCTCTTCTGCCCATTAACATCAAAAGCACCACATATTTGCTTTCAGGCAATTAATCTCAAC 979  
Db |||||||

QY 781 ATCCCAATCTTCAGTACCTGATCCATTAAGCTTTGGGGCTTCTTTCATACGACGAAGGCTC 840  
DB |||||  
QY 980 ATCCCAATCTTCAGTACCTGATCCATTAAGCTTTGGGGCTTCTTTCATACGACGAAGGCTC 1039  
DB |||||  
QY 841 GATGAACACACGATGGAACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
DB |||||  
QY 1040 GATGAACACACGATGGAACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1099  
DB |||||  
QY 901 GTTGAATTTACTTCGACAGCAGCAATCTTCGGAGATCTTCCTGGAAGGCAACGCTTCTAGG 960  
DB |||||  
QY 1100 GTTGAATTTACTTCGACAGCAGCAATCTTCGGAGATCTTCCTGGAAGGCAACGCTTCTAGG 1159  
DB |||||  
QY 961 AGTGGAAAAACCTCTTGTGCTCGGGCAGAGACTTTTGTGCACTGTGGTAGATCTCTGTCT 1020  
DB |||||  
QY 1160 AGTGGAAAAACCTCTTGTGCTCGGGCAGAGACTTTTGTGCACTGTGGTAGATCTCTGTCT 1219  
DB |||||  
QY 1021 ACCAATGTCTATCCAGACATCTTTGATTAATACCTGTTGGAATCTCTATGATCCATATC 1080  
DB |||||  
QY 1220 ACCAATGTCTATCCAGACATCTTTGATTAATACCTGTTGGAATCTCTATGATCCATATC 1279  
DB |||||  
QY 1081 GAAGTCACTACATGTTGAACAACTGGGCAAACTTAAGAAATGAGAGCCTGTGGAGT 1140  
DB |||||  
QY 1280 GAAGTCACTACATGTTGAACAACTGGGCAAACTTAAGAAATGAGAGCCTGTGGAGT 1339  
DB |||||  
QY 1141 GTAGCAAGAGGTGTTATTAGAATGTTACGAAAAAACTATGTTGTGTCGAGTGGATTTT 1200  
DB |||||  
QY 1340 GTAGCAAGAGGTGTTATTAGAATGTTACGAAAAAACTATGTTGTGTCGAGTGGATTTT 1399  
DB |||||  
QY 1201 GCACAGCAATTTCTTTAAAGGAATATTAGAAGCAAAAGTCAAGAAACCGGTGTCTGCT 1260  
DB |||||  
QY 1400 GCACAGCAATTTCTTTAAAGGAATATTAGAAGCAAAAGTCAAGAAACCGGTGTCTGCT 1459  
DB |||||  
QY 1261 CTACTTTCCCTGGAGCAGAGCTTTGTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320  
DB |||||  
QY 1460 CTACTTTCCCTGGAGCAGAGCTTTGTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1519  
DB |||||  
QY 1321 GCTGATGAAGGTAGAGACAGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
DB |||||  
QY 1520 GCTGATGAAGGTAGAGACAGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1579  
DB |||||  
QY 1381 CGAAGGAGGTGATGTCAAATCTGGCTGAGCATATTTCTATTCATCTGCTAGCAAGTCTGT 1440  
DB |||||  
QY 1580 CGAAGGAGGTGATGTCAAATCTGGCTGAGCATATTTCTATTCATCTGCTAGCAAGTCTGT 1639  
DB |||||  
QY 1441 GCCATTATGTCACACACATGTTGGCTTGGCTGCTCTCTACAGACAGGCGAGGAAT 1500  
DB |||||  
QY 1640 GCCATTATGTCACACACATGTTGGCTTGGCTGCTCTCTACAGACAGGCGAGGAAT 1699  
DB |||||  
QY 1501 GATCTCTCCACATTTGGTCGAAGACTCTTTGTGATGAAGAGGAAGTCTTGGCTCGTGAT 1560  
DB |||||  
QY 1700 GATCTCTCCACATTTGGTCGAAGACTCTTTGTGATGAAGAGGAAGTCTTGGCTCGTGAT 1759  
DB |||||  
QY 1561 TTTGACCTGGGGTCTCAGGAATTCAGAAATTCAGAAATGATAGTAATGATGCCATACAGCTGCTG 1620  
DB |||||  
QY 1760 TTTGACCTGGGGTCTCAGGAATTCAGAAATTCAGAAATGATAGTAATGATGCCATACAGCTGCTG 1819  
DB |||||  
QY 1621 GGAATTTGTCATCAATCACCACACTAGCAGGAACGATGAGTTTTTATCACCCCCAGC 1680  
DB |||||  
QY 1820 GGAATTTGTCATCAATCACCACACTAGCAGGAACGATGAGTTTTTATCACCCCCAGC 1879  
DB |||||  
QY 1681 ACAACTGTCCCATCAGTCTTTCGAATCACTTCTACAGCAATGGGTACTTCAATGCTTTT 1740  
DB |||||  
QY 1880 ACAACTGTCCCATCAGTCTTTCGAATCACTTCTACAGCAATGGGTACTTCAATGCTTTT 1939  
DB |||||  
QY 1741 ATCATGAGGCCATCATAGCTTTCAGCTTTTATGCAAGTCTGTAACAGAGGGGACTGGGG 1800  
DB |||||  
QY 1940 ATCATGAGGCCATCATAGCTTTCAGCTTTTATGCAAGTCTGTAACAGAGGGGACTGGGG 1999  
DB |||||  
QY 1801 GGTCCCACTAGACCCCACTACCTGATCAGCAGGAGCAGCTGTGGAGAGGGGCC 1860  
DB |||||  
QY 2000 GGTCCCACTAGACCCCACTACCTGATCAGCAGGAGCAGCTGTGGAGAGGGGCC 2059  
DB |||||  
QY 1861 AGCTGTGCTACCTTCTCTCCAAATGAAGGCAACCATCTCACTGCTTGGCCAGACATTTTAC 1920  
DB |||||

DB 2060 AGCTGTGCTACTCTTCTTCCAAATGAAGGCACCATCTCACTGCTTGCAGACATTTTAC 2119  
QY |||||  
DB 1921 CAAAGTCTGCCATGAACAGTAGAAGATTTATCCAGATATGGCAATCTTACAGTGGCAGAG 1980  
DB |||||  
DB 2120 CAAAGTCTGCCATGAACAGTAGAAGATTTATCCGGTATGGCAATCTTACAGTGGCAGAG 2179  
QY |||||  
DB 1981 CACGATGACGACGAGAGATATCAGTCTTCTAGTCTTGTCTGAGCAGCAGTGGGACAAAGAGCTT 2040  
DB |||||  
DB 2180 CACGATGACGACGAGAGATATCAGTCTTCTAGTCTTGTCTGAGCAGCAGTGGGACAAAGAGCTT 2239  
QY |||||  
DB 2041 CCAGAACCTTTCTCTTGGGAAAGTATGATGAAGAAAGATGAAGACAGTACTTTTGGGGAGGAA 2100  
DB |||||  
DB 2240 CCAGAACCTTTCTCTTGGGAAAGTATGATGAAGAAAGATGAAGACAGTACTTTTGGGGAGGAA 2299  
QY |||||  
DB 2101 CAGCGAGATGCTCCTCTGAGGTGAGCCCAATCCAGGAGCACCAGCAGTTTATCACCTTTC 2160  
DB |||||  
DB 2300 CAGCGAGATGCTCCTCTGAGGTGAGCCCAATCCAGGAGCACCAGCAGTTTATCACCTTTC 2359  
QY |||||  
DB 2161 TTACAGAGACTCTTGGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2220  
DB |||||  
DB 2360 TTACAGAGACTCTTGGGGCTTTGCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTTAC 2419  
QY |||||  
DB 2221 AACTTCAGTGGTCTCTTTCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAATA 2280  
DB |||||  
DB 2420 AACTTCAGTGGTCTCTTTCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAATA 2479  
QY |||||  
DB 2281 ACCAGAACAGAAAGAAATGTTCCAGTATATGCTGAGAGTGGCCACATATTTGTTGTGAAG 2340  
DB |||||  
DB 2480 ACCAGAACAGAAAGAAATGTTCCAGTATATGCTGAGAGTGGCCACATATTTGTTGTGAAG 2539  
QY |||||  
DB 2341 AATGCTGTCAAAATGTTTAAAGGATATGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTG 2400  
DB |||||  
DB 2540 AATGCTGTCAAAATGTTTAAAGGATATGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTG 2599  
QY |||||  
DB 2401 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAACAACTTCTAGAA 2460  
DB |||||  
DB 2600 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAACAACTTCTAGAA 2659  
QY |||||  
DB 2461 TATATTTCTGAGTTTTTGTGTGCTG 2484  
DB |||||  
DB 2660 TATATTTCTGAGTTTTTGTGTGCTG 2683

RESULT 7  
HSM804372  
LOCUS  
DEFINITION  
Homo sapiens mRNA; cDNA DKFp451B115 (from clone DKFp451B115);  
complete cds.  
ACCESSION  
AL833061  
VERSION  
AL833061.2 GI:30268346  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 6387)  
Bloeker, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C.,  
Fobo, G., Han, M., Osanger, A. and Wiemann, S.  
Direct Submission  
Submitted (30-APR-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuberberg, GERMANY  
COMMENT  
On Apr 30, 2003 this sequence version replaced gi:21733651.  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
This clone (DKFp451B115) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further  
information about the clone and the sequencing project is available  
at <http://mips.gsf.de/proj/cDNA/>.

FEATURES  
source

Location/Qualifiers  
1. .6387  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="10q25.2"  
/clone="DKFZp451B1115"  
/tissue\_type="CDNA-collection"  
/clone\_lib="451 (synonym: hlcel). Vector pSPORT1; host  
DH10B; sites NotI + SalI"  
/dev\_stage="adult"  
1. .6387  
/gene="DKFZp451B1115"  
197. .2683  
/gene="DKFZp451B1115"  
/note="glycerol-3-phosphate acyltransferase"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAD99967.1"  
/db\_xref="GI:30268347"  
/translation="MDESALTGLTIDVSYLPHSSEYSGVRCKHITSEWBGCFRPTVF  
RSATLWKESLMRSKRPFVGRCSYCTPQSWDKFPNPSIPSLGRNVYINETHRUR  
GWLARLSYVLFIOERDVHKMPATNTENVLNSRVQEAIAEVAALNPDGSAQOOS  
KAVNKVKKAKRILOEMVATVSPAMIRLTGWLKLENSFFWNIQHKQLEWVKAT  
ETNLPFLFPVHRSHIDYLLTFILFCHNIKAPIASGNLNIPFISTLIHKLGFFI  
RRRUDETPDGRKQVLYRALLGHIVLELRQOQFLEIFEGTRSRSRSGTSCARAGLSV  
VDTLSTNPIDILIIIVGISYGRRIIEGHYNGEQLGPKPKNESIWSVARGVIRMLRN  
YGVVRPAOPFSLNLESOSQKPSVALLSLEOALLPAILPRSPDAADDEGRDTSN  
ESRNATDESRRRLIANLEHILFTASKSCALMSTHIVACLLYVHRQGDLSLVED  
FVYKESVLARDPDLGSGNSDEDVVMHAIQLLGNCVITHTSRNDEFITPSTVPVS  
FELNYSVGLVHFIMAIACSLYVNLNKGIGGSTPPNLSISQOLVRKASLCY  
LLSNEATISLPQTFYQVCHETVGKFIQYGLITVAEHDDQDI:SPSLAEQQWKKLPE  
PLSWRDEDEDSDFGEQORDCYLVKSQKHEQOIFLQRLPLLEAYSSAAIFVH  
NFGSPVPEVLOKLKHLITRTERNVAVVAESATYCLVKNVAVMPKNDIGVFKETQK  
RVSVLELSNFFLPQCNKQLLEYILSFVL"  
6346. .6351  
/gene="DKFZp451B1115"  
polyA\_signal  
6364  
/gene="DKFZp451B1115"  
polyA\_site  
BASE COUNT 1865 a 1239 c 1381 g 1902 t  
ORIGIN

Query Match 99.7%; Score 2477.6; DB 9; Length 6387;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2480; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	1	ATGGATGAATCTGCACCTGGTGTACATAGATGTTCTTATCTGCCACATTCATCA	60
Db	197	ATGGATGAATCTGCACCTGGTGTACATAGATGTTCTTATCTGCCACATTCATCA	256
Qy	61	GAATACAGTGTGTGTCAGTGAACACACACAGTGGGAGTGGGTGAGTGGCTTTAGA	120
Db	257	GAATACAGTGTGTGTCAGTGAACACACACAGTGGGAGTGGGTGAGTGGCTTTAGA	316
Qy	121	CCACCGTCTTCAGATCTGCACTTTAAATGGAAGAAAGCCCTTAATGAGTCGGAAGG	180
Db	317	CCACCGTCTTCAGATCTGCACTTTAAATGGAAGAAAGCCCTTAATGAGTCGGAAGG	376
Qy	181	CCATTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGCTGGGACAAATTTTCAAC	240
Db	377	CCATTGTTGGAAGATGTTGTTACTCTGCACTCCCGAGCTGGGACAAATTTTCAAC	436
Qy	241	CCAGTATCCCGTCTTTGGGTTTCGGGAATGTTATTTATATCAATGAACCTCACACA	300
Db	437	CCAGTATCCCGTCTTTGGGTTTCGGGAATGTTATTTATATCAATGAACCTCACACA	496
Qy	301	CACCGCGATGGCTTGCAAGACGCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTG	360
Db	497	CACCGCGATGGCTTGCAAGACGCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTG	556
Qy	361	CATAGGGCATGTTTGGCACCAATGTGACGTGAAATGTGCTGAACACAGTAGATCA	420
Db	557	CATAGGGCATGTTTGGCACCAATGTGACGTGAAATGTGCTGAACACAGTAGATCA	616

Qy	421	GAGCAATTGCAGAAGTGGCTGTGAATTAACCCCTGATGGTTCTGCCAGCAATCA	480
Db	617	GAGCAATTGCAGAAGTGGCTGTGAATTAACCCCTGATGGTTCTGCCAGCAATCA	676
Qy	481	AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTAAAGGATCTTCAAGAAATGGTGGCACT	540
Db	677	AAAGCCGTTAAACAAAGTGAAGAAAGAAAGCTAAAGGATCTTCAAGAAATGGTGGCACT	736
Qy	541	GTCTCACCGCAATGATCAGACTGACTGGGTGGTGTCTTAAACCTGTTCAACAGCTTC	600
Db	737	GTCTCACCGCAATGATCAGACTGACTGGGTGGTGTCTTAAACCTGTTCAACAGCTTC	796
Qy	601	TTTTGGAACATTCAAATTCACAAAGGTCAACTTTCAGATGGTTAAAGCTGCAACTGAGACG	660
Db	797	TTTTGGAACATTCAAATTCACAAAGGTCAACTTTCAGATGGTTAAAGCTGCAACTGAGACG	856
Qy	661	AATTGGCGCTTCTGTCTTACAGATTCATAGATCCCATATTCATATCTGCTGCTCACT	720
Db	857	AATTGGCGCTTCTGTCTTACAGATTCATAGATCCCATATTCATATCTGCTGCTCACT	916
Qy	721	TTCAATCTCTTCTGCCATACATCAAGACCATACATCTTCCAGGCAATTAATCTCAAC	780
Db	917	TTCAATCTCTTCTGCCATACATCAAGACCATACATCTTCCAGGCAATTAATCTCAAC	976
Qy	781	ATCCCAATCTTCACTACCTTTCATAGATCCCAAGCTTGGGGGCTTCTTCATACGACGAAGCTC	840
Db	977	ATCCCAATCTTCACTACCTTTCATAGATCCCAAGCTTGGGGGCTTCTTCATACGACGAAGCTC	1036
Qy	841	GATGAAACACAGATGGAACGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA	900
Db	1037	GATGAAACACAGATGGAACGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA	1096
Qy	901	GTTGAATTTACTCGACAGCAGCAATCTTTCGAGATCTTCTGAGGACACACGTTCTTAGG	960
Db	1097	GTTGAATTTACTCGACAGCAGCAATCTTTCGAGATCTTCTGAGGACACACGTTCTTAGG	1156
Qy	961	AGTGGAAAAACCTCTTGTGCTCGGGCAGGACTTTTGTGAGTGTGGTAGATCTCTGCTCT	1020
Db	1157	AGTGGAAAAACCTCTTGTGCTCGGGCAGGACTTTTGTGAGTGTGGTAGATCTCTGCTCT	1216
Qy	1021	ACCAATGTCTATCCAGACATCTTGTGTAATAACCTTGTGGAAATCTCTATGATTCGATATC	1080
Db	1217	ACCAATGTCTATCCAGACATCTTGTGTAATAACCTTGTGGAAATCTCTATGATTCGATATC	1276
Qy	1081	GAAGTCACTACAAATGTTGAACAACTGGGCAACCTTAAGAAATGAGAGCTGTGAGT	1140
Db	1277	GAAGTCACTACAAATGTTGAACAACTGGGCAACCTTAAGAAATGAGAGCTGTGAGT	1336
Qy	1141	GTAGCAAGAGTGTATTAGAAATGTTACGAAAAAACTATGTTGTGTGTCGAGTGGATTTT	1200
Db	1337	GTAGCAAGAGTGTATTAGAAATGTTACGAAAAAACTATGTTGTGTGTCGAGTGGATTTT	1396
Qy	1201	GCACAGCAATTTTCTTAAAGGAATATTAGAAGCAAAAGTCAAGAACCCGGTGTCTGCT	1260
Db	1397	GCACAGCAATTTTCTTAAAGGAATATTAGAAGCAAAAGTCAAGAACCCGGTGTCTGCT	1456
Qy	1261	CTACTTTCCCTGGAGCAGAGGTTGTACAGCTATATCTTCTTCAAGACCCAGTGTGCT	1320
Db	1457	CTACTTTCCCTGGAGCAGAGGTTGTACAGCTATATCTTCTTCAAGACCCAGTGTGCT	1516
Qy	1321	GCTGATGAAGGTAGACACGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA	1380
Db	1517	GCTGATGAAGGTAGACACGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA	1576
Qy	1381	CGAAGGAGGTGATTTGCAAAATCTGGCTGAGCATATTTCTATTCTAGTCAAGTCTCTGT	1440
Db	1577	CGAAGGAGGTGATTTGCAAAATCTGGCTGAGCATATTTCTATTCTAGTCAAGTCTCTGT	1636
Qy	1441	GCAATTTATGTCACACATTTGGCTGTGCTGCTCTCTTACAGACACAGCGAGGAATT	1500
Db	1637	GCAATTTATGTCACACATTTGGCTGTGCTGCTCTCTTACAGACACAGCGAGGAATT	1696

```
QY 1501 GATCTCTCCACATTTGTCGAAGACTTCTTTGTGATGAAAGAGAACTCTCTGGCTCGTGAT 1560
Db 1697 GATCTCTCCACATTTGTCGAAGACTTCTTTGTGATGAAAGAGAACTCTCTGGCTCGTGAT 1756
QY 1561 TTTGACCTGGGTTCTCAGAAATTCAGAAAGATGTAGTATGATGCCATACAGCTGCTG 1620
Db 1757 TTTGACCTGGGTTCTCAGAAATTCAGAAAGATGTAGTATGATGCCATACAGCTGCTG 1816
QY 1621 GGAATTTGTGTACAAATCCACACTAGCAGGAAGATGATTTTATATACCCCCAGC 1680
Db 1817 GGAATTTGTGTACAAATCCACACTAGCAGGAAGATGATTTTATACCCCCAGC 1876
QY 1681 ACAACTGTCCCATCAGTCTTCGAATCAACTCTACAGCAATGGGTACTTCATGTCTTT 1740
Db 1877 ACAACTGTCCCATCAGTCTTCGAATCAACTCTACAGCAATGGGTACTTCATGTCTTT 1936
QY 1741 ATCATGGAGGCCATCATAGCTTGACGCTTTATGAGTTCTGAAAGAGGAGCTGGGG 1800
Db 1937 ATCATGGAGGCCATCATAGCTTGACGCTTTATGAGTTCTGAAAGAGGAGCTGGGG 1996
QY 1801 GGTCCCACTAGCACCCCACTTAACCTGATCAGCCAGGACGCTGTGGGAGGGGCC 1860
Db 1997 GGTCCCACTAGCACCCCACTTAACCTGATCAGCCAGGACGCTGTGGGAGGGGCC 2056
QY 1861 AGCCTGTGCTACCTCTCTCCAATGAAGCACCATCTCACTGCTTGCAGACATTTTAC 1920
Db 2057 AGCCTGTGCTACCTCTCTCCAATGAAGCACCATCTCACTGCTTGCAGACATTTTAC 2116
QY 1921 CAACTGTCCATGAACACAGTAGGAAAGTTATTCAGTATGGCAATCTTACAGTGGCAGAG 1980
Db 2117 CAACTGTCCATGAACACAGTAGGAAAGTTATTCAGTATGGCAATCTTACAGTGGCAGAG 2176
QY 1981 CACGATGACGAGAGATATCAGTCTGCTGCTGAGCAGCAGTGGGACAGAGCTT 2040
Db 2177 CACGATGACGAGAGATATCAGTCTGCTGCTGAGCAGCAGTGGGACAGAGCTT 2236
QY 2041 CCAGAACCTTTGTCTTGGAGAGTGTATGAAGAGATGAAGACAGTCACTTTGGGAGGAA 2100
Db 2237 CCTGAACCTTTGTCTTGGAGAGTGTATGAAGAGATGAAGACAGTCACTTTGGGAGGAA 2296
QY 2101 CAGCGAGATGCTTACCTGAAGTGAAGCAATCCAGGAGCACAGAGTTTATCACCTTC 2160
Db 2297 CAGCGAGATGCTTACCTGAAGTGAAGCAATCCAGGAGCACAGAGTTTATCACCTTC 2356
QY 2161 TTACAGAGACTCCTTGGGCTTGTCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTAC 2220
Db 2357 TTACAGAGACTCCTTGGGCTTGTCTGGAGGCTTACAGCTCTGCTGCCATCTTTGTTAC 2416
QY 2221 AACTTCAGTGGTCTCTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTAATA 2280
Db 2417 AACTTCAGTGGTCTCTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTAATA 2476
QY 2281 ACCAGAACAGAAAGTGTGAGTATATGCTGAGAGTGCCACATATTTGTTGGAAG 2340
Db 2477 ACCAGAACAGAAAGTGTGAGTATATGCTGAGAGTGCCACATATTTGTTGGAAG 2536
QY 2341 AATGCTGTCAAAATGTTAAGATATGTTGGGTTTTCAAGGAGCCAAACAAAGAGAGTG 2400
Db 2537 AATGCTGTCAAAATGTTAAGATATGTTGGGTTTTCAAGGAGCCAAACAAAGAGAGTG 2596
QY 2401 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAACCTTCTAGAA 2460
Db 2597 TCTGTTTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAACCTTCTAGAA 2656
QY 2461 TATATTCTCAGTTTTGTGGTGCTG 2484
Db 2657 TATATTCTCAGTTTTGTGGTGCTG 2680
```

## RESULT 8

HSM803772

LOCUS

DEFINITION Homo sapiens mRNA; cDNA DKF2p313H071 (from clone DKF2p313H071).

5689 bp

mRNA

linear

PRI 13-MAY-2003

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISMAL832464  
AL832464.1 GI:21733030

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 5689)

Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.

Direct Submission

Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKFZp313H071) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/.

## FEATURES

Source

1..5689

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="DKFZp313H071"

/tissue\_type="human fetal skin"

/clone\_lib="313 (synonym: hlcc2). Vector pTriplex2; host

DH10B; sites SfiIA + SfiIB"

/dev\_stage="fetal"

5654

polyA\_site

BASE COUNT 1521 a 1149 c 1256 g 1763 t

## ORIGIN

Query Match 85.0%; Score 2111.8; DB 9; Length 5689;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2116; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGGATGATCTGCACTGACCCCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 60

Db 167 ATGGATGATCTGCACTGACCCCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 226

QY 61 GAATACAGTGTGGTTCGATGTAAGCACACAAAGTGAAGATGGGGTGAAGTGGCTTTAGA 120

Db 227 GAATACAGTGTGGTTCGATGTAAGCACACAAAGTGAAGATGGGGTGAAGTGGCTTTAGA 286

QY 121 CCACCGTCTTCAGATCTGCACTTAAATGGAAGAAAGCCTAATGAGTCGAAAGG 180

Db 287 CCACCGTCTTCAGATCTGCACTTAAATGGAAGAAAGCCTAATGAGTCGAAAGG 346

QY 181 CCATTGTTGGAAGATGTTGTTACTCTCCAGAGCTCCCGACAGCAAAATTTTCAAC 240

Db 347 CCATTGTTGGAAGATGTTGTTACTCTCCAGAGCTCCCGACAGCAAAATTTTCAAC 406

QY 241 CCAGTATCCCGTCTTTGGGTTTGGGAATGTTATATCAATGAAACTCACACAAG 300

Db 407 CCAGTATCCCGTCTTTGGGTTTGGGAATGTTATATCAATGAAACTCACACAAG 466

QY 301 CACCGGATGGCTTGCAGAGCCCTTCTTACGTTCTTTTATTCAGAGCAGATGG 360

Db 467 CACCGGATGGCTTGCAGAGCCCTTCTTACGTTCTTTTATTCAGAGCAGATGG 526

QY 361 CATAAGGGCATGTTTGGCCACCAATGTGACTGAAAAATGCTGAAAGAGTAGATACAA 420

Db 527 CATAAGGGCATGTTTGGCCACCAATGTGACTGAAAAATGCTGAAAGAGTAGATACAA 586

QY 421 GAGGCAATTCAGAGTGGCTGCTGAAATTAACCCCTGATGTTCTGCCAGCAGCAATCA 480

Db 587 GAGGCAATTCAGAGTGGCTGCTGAAATTAACCCCTGATGTTCTGCCAGCAGCAATCA 646



Qy 481 AAAGCGTTAA CAAAGTGAAGAAAGCTAAAGGATTTCTTCAAGAAATGGTTGCCACT 540  
Db |||||  
647 AAAGCGTTAA CAAAGTGAAGAAAGCTAAAGGATTTCTTCAAGAAATGGTTGCCACT 706  
Qy 541 GTCTCACCAGCAATGATCAGACTGACTGGTGGTGGTCTGCTAAACTGTTCAACAGCTTC 600  
Db |||||  
707 GTCTCACCAGCAATGATCAGACTGACTGGTGGTGGTCTGCTAAACTGTTCAACAGCTTC 766  
Qy 601 TTTTGGAACTTCAAATTCACAAAGGTCAA CTTGAGATGGTTAAAGCTGCAACTGAGACG 660  
Db |||||  
767 TTTTGGAACTTCAAATTCACAAAGGTCAA CTTGAGATGGTTAAAGCTGCAACTGAGACG 826  
Qy 661 AATTGGCCGCTTCTGTTTCTACAGTTCAATAGATCCCATATGACTATCTGCTGCTCACT 720  
Db |||||  
827 AATTGGCCGCTTCTGTTTCTACAGTTCAATAGATCCCATATGACTATCTGCTGCTCACT 886  
Qy 721 TTCAATCTCTTCTGTCATTAACATCAAGACCAATACATTTGCTTCAAGCAATTAATCTCAAC 780  
Db |||||  
887 TTCAATCTCTTCTGTCATTAACATCAAGACCAATACATTTGCTTCAAGCAATTAATCTCAAC 946  
Qy 781 ATCCCAATCTTTCAGTACCTTGCATTAAGCTTGGGGCTTCTTCAACAGCAGAGGCTC 840  
Db |||||  
947 ATCCCAATCTTTCAGTACCTTGCATTAAGCTTGGGGCTTCTTCAACAGCAGAGGCTC 1006  
Qy 841 GATGAAACAC CAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db |||||  
1007 GATGAAACAC CAGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1066  
Qy 901 GTTGAATTTACTTGACAGCAGCAATCTTGAGATCTTCTGGAAGGACACGTTCTAGG 960  
Db |||||  
1067 GTTGAATTTACTTGACAGCAGCAATCTTCTGGAAGGACACGTTCTAGG 1126  
Qy 961 AGTGGAAGAACTCTTGCTCGGCGAGACTTTTGTCAGTTGTGTAGTACTCTGTCT 1020  
Db |||||  
1127 AGTGGAAGAACTCTTGCTCGGCGAGACTTTTGTCAGTTGTGTAGTACTCTGTCT 1186  
Qy 1021 ACCAATGTATCCCGACATCTTGATATACCTGTGGAATCTCCTATGATCGCATATC 1080  
Db |||||  
1187 ACCAATGTATCCCGACATCTTGATATACCTGTGGAATCTCCTATGATCGCATATC 1246  
Qy 1081 GAAGGTCTACTCAATGTTGAACAACTGGGAAACCTTAAGAAATGAGAGCTGTGGAGT 1140  
Db |||||  
1247 GAAGGTCTACTCAATGTTGAACAACTGGGAAACCTTAAGAAATGAGAGCTGTGGAGT 1306  
Qy 1141 GTAGCAGAGCTGTTATTAGATGTTACGAAAACATGTTGTTGTCGAGTGGATTTT 1200  
Db |||||  
1307 GTAGCAGAGCTGTTATTAGATGTTACGAAAACATGTTGTTGTCGAGTGGATTTT 1366  
Qy 1201 GCACAGCCATTTCTTAAAGGAATATTAGAAAGCCAAAGTCAAGAACCCGGTGTCTGCT 1260  
Db |||||  
1367 GCACAGCCATTTCTTAAAGGAATATTAGAAAGCCAAAGTCAAGAACCCGGTGTCTGCT 1426  
Qy 1261 CTACTTTCCCTGGAGCAAGCTGTTTACAGTATATCTTCTTCAAGACCCAGTGTGCT 1320  
Db |||||  
1427 CTACTTTCCCTGGAGCAAGCTGTTTACAGTATATCTTCTTCAAGACCCAGTGTGCT 1486  
Qy 1321 GCTGATGAAGTGTAGACACCTGCTTAAATGATCCAGAAATGCAACAGATCAATCCCTA 1380  
Db |||||  
1487 GCTGATGAAGTGTAGACACCTGCTTAAATGATCCAGAAATGCAACAGATCAATCCCTA 1546  
Qy 1381 CGAAGGAGGTTGATTGCAAACTGCTGCTGAGCATATTTCTATTCTAGCAAGTCTCTGT 1440  
Db |||||  
1547 CGAAGGAGGTTGATTGCAAACTGCTGCTGAGCATATTTCTATTCTAGCAAGTCTCTGT 1606  
Qy 1441 GCCATTATGTCCACACATTTGGCTGCTGCTCTCTCTACAGACACAGCAGGGAATT 1500  
Db |||||  
1607 GCCATTATGTCCACACATTTGGCTGCTGCTCTCTCTCTACAGACACAGCAGGGAATT 1666  
Qy 1501 GATCTCTCCATTTGGTGTGAGACTTCTTTGTGATGAAGGAAGTCTCTGGCTGTGAT 1560  
Db |||||  
1667 GATCTCTCCATTTGGTGTGAGACTTCTTTTGTGATGAAGGAAGTCTCTGGCTGTGAT 1726

Qy 1561 TTTGACCTGGGTTCTTCAGGAAATTCAGAAATAGTAGTAAATGATGCATGCATACAGTGTCTG 1620  
Db |||||  
1727 TTTGACCTGGGTTCTTCAGGAAATTCAGAAATAGTAGTAAATGATGCATGCATACAGTGTCTG 1786  
Qy 1621 GGAATTTGTGTCAAAATTCACCCACACTAGCAGGAACGATGATGTTTTTATCACCCCAAGC 1680  
Db |||||  
1787 GGAATTTGTGTCAAAATTCACCCACACTAGCAGGAACGATGATGTTTTTATCACCCCAAGC 1846  
Qy 1681 ACAACTGTCCCATCAGTCTTCGAACTCTTCTACAGCAATGGGTACTTCTGTCTTTT 1740  
Db |||||  
1847 ACAACTGTCCCATCAGTCTTCGAACTCTTCTACAGCAATGGGTACTTCTGTCTTTT 1906  
Qy 1741 ATCATGAGGCAATCATAGCTTTCAGCTTTATGCAAGTTCTGAACAAGAGGGACTTGGGG 1800  
Db |||||  
1907 ATCATGAGGCAATCATAGCTTTCAGCTTTTATGCAAGTTCTGAACAAGAGGGACTTGGGG 1966  
Qy 1801 GGTCCCACTAGCACCCCACTTAACCTGATCAGCAGCAGCAGCTGGTGGGAAGGCGGCC 1860  
Db |||||  
1967 GGTCCCACTAGCACCCCACTTAACCTGATCAGCAGCAGCAGCTGGTGGGAAGGCGGCC 2026  
Qy 1861 AGCTGTGTACCTTCTCTCAATGAAGGCAACCATCTCACTGCCTTGGCAGACATTTTAC 1920  
Db |||||  
2027 AGCTGTGTACCTTCTCTCAATGAAGGCAACCATCTCACTGCCTTGGCAGATTTTAC 2086  
Qy 1921 CAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTCTTACAGTGGCAGAG 1980  
Db |||||  
2087 CAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTCTTACAGTGGCAGAG 2146  
Qy 1981 CACGATCACCGAAGATATCAGTCTTCTGCTGAGCAGAGTGGGCAAGAAGCTT 2040  
Db |||||  
2147 CACGATCACCGAAGATATCAGTCTTCTGCTGAGCAGAGTGGGCAAGAAGCTT 2206  
Qy 2041 CCAGAACTTTGCTTCTGAGAAAGTATGAAGACATGAAGACATGACTTTTGGGGAGGAA 2100  
Db |||||  
2207 CCTGAACCTTTGCTTCTGAGAAAGTATGAAGACATGAAGACATGACTTTTGGGGAGGAA 2266  
Qy 2101 CAGCGAGATTGCTACCTGAAGGT 2123  
Db |||||  
2267 CAGCGAGATTGCTACCTGAAGGT 2289

## RESULT 9

LOCUS AB046780 5674 bp mRNA linear PRI 22-FEB-2001  
DEFINITION Homo sapiens mRNA for KIAA1560 protein, partial cds.  
ACCESSION AB046780  
VERSION AB046780.1 GI:10047184  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Nagase, T., Kikuno, R., Nakayama, M., Hirose, M. and Ohara, O.  
TITLE Prediction of the coding sequences of unidentified human genes.  
XVIII. The complete sequences of 100 new cDNA clones from brain  
which code for large proteins in vitro  
JOURNAL DNA Res. 7 (4), 273-281 (2000)  
MEDLINE 20450683  
PUBMED 10997877  
REFERENCE 2 (bases 1 to 5674)  
AUTHORS Ohara, O., Nagase, T. and Kikuno, R.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute,  
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba  
292-0812, Japan (E-mail:cdmainfo@kazusa.or.jp,  
URL: http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,  
Fax:81-438-52-3914)  
FEATURES  
Location/Qualifiers  
1..5674  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"



		/clone="fh19959"		/notes="vector:pBluescriptII SK plus"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
--	--	------------------	--	---------------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

QY	1159	AGAAATGTTACGAAAAAATATGTTGTGTCGAGTGGATTTTGCACAGCAATTTTCTTTA	1218
DB	661	AGAAATGTTACGAAAAAATATGTTGTGTCGAGTGGATTTTGCACAGCAATTTTCTTTA	720
QY	1219	AAGGAATATTTAGAAAGCCAAAGTCAAGAACCGGTGTCTGCTCTACTTTCCCTGGAGCAA	1278
DB	721	AAGGAATATTTAGAAAGCCAAAGTCAAGAACCGGTGTCTGCTCTACTTTCCCTGGAGCAA	780
QY	1279	GGTTGTGTACCAAGCTATCTCTCTCAAGACCCAGTGATGCTGCTGATGAAGTGTAGAC	1338
DB	781	GGTTGTGTACCAAGCTATCTCTCTCAAGACCCAGTGATGCTGCTGATGAAGTGTAGAC	840
QY	1339	ACGTCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGAGGTTGATTGCA	1398
DB	841	ACGTCATTAAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGAGGTTGATTGCA	900
QY	1399	AATCTGGCTGAGCATATTTCTATCTACTGTAGCAAGTCTCTGTGCCATTTATGCCACAC	1458
DB	901	AATCTGGCTGAGCATATTTCTATCTACTGTAGCAAGTCTCTGTGCCATTTATGCCACAC	960
QY	1459	ATTGTGGCTTGGCTGCTCTCTACAGACACGACGAGGAATTTGATCTCTCCACATTTGTC	1518
DB	961	ATTGTGGCTTGGCTGCTCTCTACAGACACGACGAGGAATTTGATCTCTCCACATTTGTC	1020
QY	1519	GAAGACTTCTTTGTGATGAAAGAGGAGTCTCTGGCTCGTGTGATTTTGACCTGGGGTCTCA	1578
DB	1021	GAAGACTTCTTTGTGATGAAAGAGGAGTCTCTGGCTCGTGTGATTTTGACCTGGGGTCTCA	1080
QY	1579	GGAAATTCAGAAAGATGTAGTAATGCAATACAGCTGCTGGGAAATTTGTGTCAATTC	1638
DB	1081	GGAAATTCAGAAAGATGTAGTAATGCAATACAGCTGCTGGGAAATTTGTGTCAATTC	1140
QY	1639	ACCCACACTAGCAGGAACGATGAGTCTTTTATACCCCGCAGCACACTCTCCATCAGTC	1698
DB	1141	ACCCACACTAGCAGGAACGATGAGTCTTTTATACCCCGCAGCACACTCTCCATCAGTC	1200
QY	1699	TTTCGAATCAACTTTCTACAGCAATGGGTGATCTTATGTCATGATGAGGCAATCATA	1758
DB	1201	TTTCGAATCAACTTTCTACAGCAATGGGTGATCTTATGTCATGATGAGGCAATCATA	1260
QY	1759	GTTGAGAGCTTTATGCGATTTCTGAACAGAGGGGATGCGGGGGTCCCACTAGCACCCCA	1818
DB	1261	GTTGAGAGCTTTATGCGATTTCTGAACAGAGGGGATGCGGGGGTCCCACTAGCACCCCA	1320
QY	1819	CCTAACCTGATCAGCCAGGAGCAGCTGGTGCGGAAGGGCGGCGGCTGTGCTTCTCTC	1878
DB	1321	CCTAACCTGATCAGCCAGGAGCAGCTGGTGCGGAAGGGCGGCGGCTGTGCTTCTCTC	1380
QY	1879	TCCAATGAAGGCACCATCTCTGCTTCCAGACATTTTACCAAGTCTGCCATGAAACA	1938
DB	1381	TCCAATGAAGGCACCATCTCTGCTTCCAGACATTTTACCAAGTCTGCCATGAAACA	1440
QY	1939	GTAGGAAGTTTATCCAGTATGCGATTTCTTACAGTGGCAGACAGATGACAGGAAGAT	1998
DB	1441	GTAGGAAGTTTATCCAGTATGCGATTTCTTACAGTGGCAGACAGATGACAGGAAGAT	1500
QY	1999	ATCAGTCTCTAGTCTTGTGCTGAGCAGCAGTGGGACAGAAAGCTTCCAGAACTTTGCTCG	2058
DB	1501	ATCAGTCTCTAGTCTTGTGCTGAGCAGCAGTGGGACAGAAAGCTTCCAGAACTTTGCTCG	1560
QY	2059	AGAAGTGTATGAAGAAGATGAAGACAGTACTTTTGGGAGGAAACAGCGAGATTTGCTTAC	2118
DB	1561	AGAAGTGTATGAAGAAGATGAAGACAGTACTTTTGGGAGGAAACAGCGAGATTTGCTTAC	1620
QY	2119	AAGGTGAGCCAAATCCAAGGAGCACCGAGAGTTTATCACTTTTACAGAGATCTCTTGGG	2178
DB	1621	AAGGTGAGCCAAATCCAAGGAGCACCGAGAGTTTATCACTTTTACAGAGATCTCTTGGG	1680
QY	2179	CTTTTGTGGAGGCTTACAGCTCTGCTGCAATCTTTTGTTCACAACTTCAAGTGTCTGTT	2238
DB	1681	CTTTTGTGGAGGCTTACAGCTCTGCTGCAATCTTTTGTTCACAACTTCAAGTGTCTGTT	1740

QY 2239 CCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTAATAACAGAACAGAAAGAAAT 2298  
Db |||||  
1741 CCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTAATAACAGAACAGAAAGAAAT 1800  
QY |||||  
2299 GTTGCAATATGCTGAGAGTGCCACATATGCTGTTGGAAGATGCTGCGAATAATGTTT 2358  
Db |||||  
1801 GTTGCAATATGCTGAGAGTGCCACATATGCTGTTGGAAGATGCTGCGAATAATGTTT 1860  
QY |||||  
2359 AAGGATATGCGGTTTCAAGGAGACCAAAACAAAGAGAGTGCTGTTTGTAGAACTGAGC 2418  
Db |||||  
1861 AAGGATATGCGGTTTCAAGGAGACCAAAACAAAGAGAGTGCTGTTTGTAGAACTGAGC 1920  
QY |||||  
2419 AGCACTTTTACCTCAATGCAACCGACAAACCTTCTAGAAATATATCTGAGTTTGTG 2478  
Db |||||  
1921 AGCACTTTTACCTCAATGCAACCGACAAACCTTCTAGAAATATATCTGAGTTTGTG 1980  
QY |||||  
2479 GTGCTG 2484  
Db |||||  
1981 GTGCTG 1986

RESULT 10  
AX401972  
LOCUS AX401972 2646 bp DNA linear PAT 06-JUN-2002  
DEFINITION Sequence 1648 from Patent WO0210453.  
ACCESSION AX401972  
VERSION AX401972.1 GI:21338152  
KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1  
AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and  
Elashoff, M.R.  
TITLE Molecular toxicology modeling  
JOURNAL Patent: WO 0210453-A 1648 07-FEB-2002;  
Gene Logic, Inc. (US)  
FEATURES  
source  
1..2646  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/note="EMBL/GenBank Accession No. NM\_017274"

BASE COUNT 670 a 689 c 670 g 617 t  
ORIGIN

Query Match 75.7%; Score 1879.2; DB 6; Length 2646;  
Best Local Similarity 84.8%; Pred. No. 0;  
Matches 2106; Conservative 0; Mismatches 378; Indels 0; Gaps 0;

QY 1 ATGATGATCTGCACTGACCTGTTGTTACATAGATGTTTCTATCTGCCACATTCATCA 60  
Db |||||  
160 ATGGAGGAGTCTTCAGTGACAAATTTGGCAATAGACGTTTCTATCTGCCAAATTCATCA 219  
QY |||||  
61 GAATACAGTGTGTCGATGTAAGCACACAGTGGGATGCGGTGAGTGGCTTTAGA 120  
Db |||||  
220 GAATACAGCTTGGCCGATGTAACACAGATGAGGACTGGGTGAGTGGCTTTCAA 279  
QY |||||  
121 CCCACCGCTTTCAGATCTGCACTTTTAAATGGAAGAAAGCCTAATGAGTCGGAAGAG 180  
Db |||||  
280 CTTACCTTTCAGATCGCAACGCTGAAATGGAAGAGACCTCATGAGCCGGAAGAG 339  
QY |||||  
181 CCAATTTGTGAAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGCAAAATTTTCAAC 240  
Db |||||  
340 CCTTCTGTTGGGAAGTGTGCTATTCATGCAAGCTCAGAGCTGGGAAAGTTTTCAC 399  
QY |||||  
241 CCCAGTATCCCGTCTTTGGGTTTGGGATGTTTATTTATATCAATGAACTCACAAGA 300  
Db |||||  
400 CCCAGTATCCCATCTCTGGGTTTGGGAATGTTTATTTATATCAATGAACTCACAAG 459  
QY |||||  
301 CACCGCGATGGCTTGCAAGACGCTTCTTACGTTCTTTTATTTCAAGACGAGATGTG 360

Db |||||  
460 CACCGAGATGGCTGGCAAGACGCTTCTTACATCCTTTTGTTCAGAGCGCATGTC 519  
QY |||||  
361 CATAAGGGCATGTTTGGCCCAATGACCTGAAATCTGTGACACAGCTAGAGTACAA 420  
Db |||||  
520 CACAAGGGCATGTTTGGCCCAATGACCTGAAATCTGTGACACAGCTAGAGTACAA 579  
QY |||||  
421 GAGCAATTCAGAAAGTGGCTGCTGAAATTAACCCCTGATGGTTCTGCCAGCAGCAATCA 480  
Db |||||  
580 GAGCAATTCGTAAGTGGCTGCAAGATTAACCCCGATGGAATCTGCCAGCAGCTCC 639  
QY |||||  
481 AAAGCCGTTTAAAGTGAAGAAAGAAAGCTTAAAGGATTTCTTCAAGAAATGTTTGCAC 540  
Db |||||  
640 AAAGCCATCCAGAAAGTGAAGAAAGAAAGCTTAAAGGATTTCTTCAAGAAATGTTTGC 699  
QY |||||  
541 GTCTCAGCGCAATGATCAGACTGACTGGGTGGGTGCTTAAACTGTTTCAACAGCTTC 600  
Db |||||  
700 GTCTCCCCCGGATGATCAGGCTGACTGGGTGGGTGTTTAAAGCTTCTTCAACAGCTTC 759  
QY |||||  
601 TTTTGAACATTCAAATTCACAAAGGTCAACTTCAGATGGTTAAAGCTGCAACTGAGACG 660  
Db |||||  
760 TTTTGAACATTCAGATTCACAAAGGTCAACTTCAGATGGTTAAAGCTGCAACTGAGACG 819  
QY |||||  
661 AATTGCGCTTCTGTTTCTTACAGTTTCTATGATCCCATATTTGATATCTGCTCACT 720  
Db |||||  
820 AATCTGCGCTTCTGTTTCTTACAGTTTCTATGATCCCATATTTGATATCTGCTCACT 879  
QY |||||  
721 TTTTCTCTCTGCTGCTTACAGTTTCTATGATCCCATATTTGATATCTGCTCACT 780  
Db |||||  
880 TTTTCTCTCTGCTGCTTACAGTTTCTATGATCCCATATTTGATATCTGCTCACT 939  
QY |||||  
781 ATCCCAATCTTCAAGTACCTTTGATTCATGATGGGTGGGTGCTTCTTCAACAGCAAGGCTC 840  
Db |||||  
940 ATCCCAATCTTCAAGTACCTTTGATTCATGATGGGTGGGTGCTTCTTCAACAGCAAGGCTC 999  
QY |||||  
841 GATGAAACACAGATGCAAGCAAGATGTTTCTTATAGAGCTTTGCTTCAATGGGCATATA 900  
Db |||||  
1000 GATGAAACACAGATGCAAGCAAGATGTTTCTTATAGAGCTTTGCTTCAATGGGCATATA 1059  
QY |||||  
901 GTTGAATTTCTTCAAGTACCTTTGATTCATGATGGGTGGGTGCTTCTTCAACAGCAAGGCTC 960  
Db |||||  
1060 GTTGAATTTCTTCAAGTACCTTTGATTCATGATGGGTGGGTGCTTCTTCAACAGCAAGGCTC 1119  
QY |||||  
961 AGTGAAGAAACCTTCTGCTGCTGCGGACGAGCTTTTGTGAGTGTGTTGATGTTCTGCT 1020  
Db |||||  
1120 AGTGAAGAAACCTTCTGCTGCTGCGGACGAGCTTTTGTGAGTGTGTTGATGTTCTGCT 1179  
QY |||||  
1021 ACCAATGTCATCCAGACATCTTTGATTAATACCTTTTGAATATCTCTTATGATCGCATATC 1080  
Db |||||  
1180 TCCAACACCATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGAT 1239  
QY |||||  
1081 GAAGGTCACTACAAATGTTGAACAACTGGGCAAACTTAAAGAAATGAGAGCTGTGAGT 1140  
Db |||||  
1240 GAAGGTCACTACAAATGTTGAACAACTGGGCAAACTTAAAGAAATGAGAGCTGTGAGT 1299  
QY |||||  
1141 GTAGCAAGGTGTTTATGATGTTTACGAAATATGTTTGTGTTGTTGTTGTTGTTGTTGTT 1200  
Db |||||  
1300 GTGCAAGAGCGTTTATCAGATGCTGCGGAAATATGCTGCTGCTGCTGCTGCTGCTGCT 1359  
QY |||||  
1201 GCACAGCATTTTCTTAAAGGAATATTTAAGAAAGCCAAAGTTCAGAAACCGGCTGCTGCT 1260  
Db |||||  
1360 GCACAGCATTTTCTTAAAGGAATATTTAAGAAAGCCAAAGTTCAGAAACCTGATCTGCT 1419  
QY |||||  
1261 CTACTTTTCTGAGCAAGCGTTGTTTACAGCTATATCTTCTTCAAGACCGCAGTGTGCT 1320  
Db |||||  
1420 CCTCTCTTTTGGAGCAAGCACTGTTTACAGCAATCTTCTTCAAGACCTGATGCTGCT 1479  
QY |||||  
1321 GCTGATGAAGGTAGAGACAGCTCCATTAATGAGTTCAGAAATGCAACAGATGATCTCTCA 1380  
Db |||||  
1480 GCTGCGCAACATGAAGACATGTTCCAGTAATGAGTTCAGAAATGCAACAGATGATCTCTCA 1539  
QY |||||  
1381 CGAAGGAGGTGTTGTTGCAATCTGGCTGAGCATATTTCTTCTGCTGCTGCTGCTGCTGCT 1440

Db 1540 CGAAGGAGGCTGATCGCAAAACCTGGGGAGACATTTCTTACCAGCAGAGTCTCTGC 1599  
Qy 1441 GCCATTATGTCACACACATGTTGGCTTGCCTCTCTACAGACACAGGAGGAAAT 1500  
Db 1600 GCTATCATGTCACACATGTTGGCTTGCCTCTCTACAGACACAGGAGGAAATC 1659  
Qy 1501 GATCTCTCCACATGTTGTCGAAGACTTCTTTGTGATGAAAGAGAAAGTCTCTGCTGAT 1560  
Db 1660 CACCTCTCCAGCTGTTGGAAGACTTCTTTGTGATGAAAGAGAAAGTCTAGTCTGGAT 1719  
Qy 1561 TTTGACCTGGGGTCTCAGAAATTCAGAAATGATGATGATGATGATGATGATGATGATG 1620  
Db 1720 TTTGACCTGGGGTCTCAGAAATTCAGAAATGATGATGATGATGATGATGATGATGATG 1779  
Qy 1621 GGAATTTGTGTCAATCAACCCACCTAGCAGAAAGATGATGATGATGATGATGATGATG 1680  
Db 1780 GGAATTTGTGTCAATCAACCCACCTAGCAGAAAGATGATGATGATGATGATGATGATG 1839  
Qy 1681 ACAACTGTCCCATCAGTCTTTCGAACCTTACAGCAATGGGTACTTTCATGCTTT 1740  
Db 1840 ACAACTGTCCCGTCCGCTTTGAACCTTACAGCAATGGGTACTTTCATGCTTT 1899  
Qy 1741 ATCATGAGGCTCATAGTCTTTCAGCTTTTATGAGTCTTTCAGCAATGGGTACTTTCATG 1800  
Db 1900 ATCATGAGGCTCATAGTCTTTCAGCTTTTATGAGTCTTTCAGCAATGGGTACTTTCATG 1959  
Qy 1801 GGTCCCACTAGCAGCCACCTTACCTGATCAGCAGGAGCAGTGTGGAGGGCGCC 1860  
Db 1960 GGTCTGCGGAGGCTTGGCAACCTGATCAGCAGGAGCAGTGTGGAGGGCGCC 2019  
Qy 1861 AGCTGTGCTACCTTCTCTCAATGAGGACCATCTCTACCTGCTTTCAGCAATGGGTACTT 1920  
Db 2020 AGCTGTGCTACCTTCTCTCAATGAGGACCATCTCTACCTGCTTTCAGCAATGGGTACTT 2079  
Qy 1921 CAAGTCTGCCATGAACAGTAGGAAGTTTATCAGTATGGCATTTTACAGTGGCAGAG 1980  
Db 2080 CAGGTTGTCAAGAGACAGTAGGAAGTTTATCAGTATGGCATTTTACAGTGGCAGAG 2139  
Qy 1981 CAGGATGACAGGAGATACAGTCTTCTGCTGAGCAGCAGTGGGACAGAGCTT 2040  
Db 2140 CAAGATGACAGGAGATGAGTCTTCTGCTGAGCAGCAGTGGGACAGAGCTT 2199  
Qy 2041 CCAGAACTTTGTCTTGGAGAGTGTGAGAGAGTGAAGACAGTGAAGTCTTGGGAGGAA 2100  
Db 2200 CCGAGGCTCTGAACCTGGAAGTGTGAGAGAGTGAAGACAGTGAAGTCTTGGGAGGAG 2259  
Qy 2101 CAGCAGATGCTTACCTGAAGTGTGAGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2160  
Db 2260 CAGCAGATGCTTACCTGAAGTGTGAGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2319  
Qy 2161 TTACAGAGCTCTTGGGCTTGTGGAGGCTTACAGTCTTCTGCTGCTGCTGCTGCTGCTG 2220  
Db 2320 CTGACAGAGCTTCTGGGCTTGTGGAGGCTTACAGTCTTCTGCTGCTGCTGCTGCTGCTG 2379  
Qy 2221 AACTTCAGTGTGCTTCTGAGAACTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2280  
Db 2380 ACCTTCGCGGCCAGTCTCGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 2439  
Qy 2281 ACCAGAACAGAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2340  
Db 2440 ACCAGAACAGAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2499  
Qy 2341 AATGCTGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2400  
Db 2500 AATGCTGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2559  
Qy 2401 TCTGTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2460  
Db 2560 TCTGTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2619  
Qy 2461 TATATCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2484  
Db 2620 TACATCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2643

## RESULT 11

AF021348

## LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .2646

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/tissue\_type="liver"

1. .2646

/gene="GPAT"

160. .2646

/gene="GPAT"

/EC\_numbers="2.3.1.15"

/note="similar to protein encoded by Mus musculus glycerol

3-phosphate acyltransferase gene, GenBank Accession Number

W7003, and Accession Numbers U36771, U36772, U36773"

/codon\_start=1

/product="sn-glycerol-3-phosphate acyltransferase"

/protein\_id="AA071605.1"

/db\_xref="GI:2444459"

/translation="MESSVTITGIDVSLPNSBSYSLGRKHTNEDWDCGKPTFF

RSATLKESLMSRKRPFVRCYSCYSPQSWERFFNPSTPSLGRNVIYNETHRR

GWLARLSVILFVQERDVHGMFATITDNVLSNRVOEIAEVAELNIPDGAQOOS

KAIQKVKARKILOEMVATVSPGMRLTGWLLKLFNSFFWNIQIHKGQLEWKAAT

ETNLELFPVHRSHIDYLLTFLFLCHNIKAPYASNNLNIPFSTLHKLGFFI

RRRDETDPKDIILYRALLGHIVELLRQQQFLFLEGTGRSGKTSKARAGLLSV

VDTLSNTIPDILVPGISYDRIEIGHYGEQIGKPKNESLSMVGAVIRMLRKN

YGVYVDFAQFSLKEYLEGQKQVSPAPLSLEQALLPAILRSPDAAAHEHDSN

BSRNAADAFRRRLIANLAHELITASKSCAISMTHIVACLLLYRHRQIHLSTLVD

FEVMEKVLARDPDLGFGNSDVVMHAIOLLGNCVTITHTSRKDEFFTPSTTPSV

FELNYSNGVLVFMIEALICSIVQNKRGSGSAGGLNLISQQLVRAASLCY

LUSNGTSLPCTFYQVQCTETGKPIQIGLITVAEQDDQEDVSPGLAEQKKWLPE

PLNWSDEDESDSDFGRQCYLKSAKHEQQTFLQRLGLGLLEAYSAATFVH

TFRGVPSEYLLQHYLLTRTENVAVYASATYCLVKNVAKMFKDGVKPKTKQK

RASVLELSTFLPQSGRQKLLLEYILSFVVL"

BASE COUNT 670 a 689 c 670 g 617 t

ORIGIN

Query Match

Best Local Similarity 84.8%; Pred. No. 0;

Matches 2106; Conservative 0; Mismatches 378; Indels 0; Gaps 0;

Qy 1 ATGGATGAATCTGCACTGACCTTGGTACAATAGATGTTTCTTATCTGCCCATTCATCA 60

Db 160 ATGGAGGAGTCTTCAGTGAATGGCAATAGACGTTTCTTATCTGCCCATTCATCA 219

QY 61 GAATACAGTGTGGTCGATGAAGCACACAACTGAGGAATGGGTGAGTGGCTTTAGA 120  
Db 220 GAATACAGCTTTGGCGGATGTAAACACACAGATGAGGACTGGGTGACTGGCTTTCAA 279  
QY 121 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCTAATCAGTCGGAAGG 180  
Db 280 CCTACTTCTTCAGATCCGAACGCTGAATGGAAGGAGGCTCATGAGCCGGAAGAGG 339  
QY 181 CCAATTGTGTGAAGATGTGTACTCTTGCACTCCCCAGAGCTGGGCAAAATTTTTCAAC 240  
Db 340 CCCTTTCGTGGGAAGGTGTGTCTATTATGTCAGCGCTCAGAGCTGGGAAGGTTTTCAAC 399  
QY 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTTATATCAATGAATCAACACACAAGA 300  
Db 400 CCCAGTATCCCATCTCTGGGTTCGGAATGTATTTATCAATGAATCAACACACAAGG 459  
QY 301 CACCCGGATGGCTTGCAGACGCTTCTTCTTACGTTCTTTTATTCAAGAGCGAGATGTG 360  
Db 460 CACCGAGGATGGCTGGCAAGACGGCTTTCTTACATCTTTTGTTCGAAGCGCGATGTC 519  
QY 361 CATAAAGGCAATGTTTGCACCAATGTGACTGAAATGTGCTGAAACAGCAGTAGAGTACAA 420  
Db 520 CACAAGGCAATGTTTGCACCACTGATCACTGACATGTACTGAATAGCAGCAGTCCAA 579  
QY 421 GAGGCAATTCGAAGTGGCTGCTGAATTAACCCCTGATGTTCTGCCCAGCAGCAATCA 480  
Db 580 GAGGCAATTCGTAAGTGGCTGCAAAATGAACCCGGATGGATCTGCCAGCAGCAGTCC 639  
QY 481 AAAGCCGTTAAACAAGTGAAGAAAGAAAGCTAAAGGATCTTCAAGAAATGGTTGCCACT 540  
Db 640 AAAGCCATCCAGAAAGTGAAGAAAGCAAGGAGATCCCTCCAGAAATGGTTGCTACA 699  
QY 541 GTCTCACCGGCAATGATCAGACTGACTGGGTGGTCTGCTAAACTGTTCAAACAGCTTC 600  
Db 700 GTCTCCCCGGGATGATCAGGCTGACTGGCTGGGTGTTACTAAAGCTCTTCAACAGCTTC 759  
QY 601 TTTTGGAACTTCAAAATTCAAAAGGTCAAACTTGAGATGGTTAAAGCTGCAACTGAGACG 660  
Db 760 TTCTGGAACTTCAAGATTCAAAGGTCAACTTGAGATGGTGAAGCTGCAACTGAGACG 819  
QY 661 AATTGGCCGCTCTGTTTCTACCACTTATAGATCCCATATGACTATCTGCTGCTACT 720  
Db 820 AATCTGCCGCTCTTGTCTGCGGTGACAGATCCCATCGACTACTCTGCTGCTACC 879  
QY 721 TTCTATCTCTTCTGCAATAATCAAGCAACATACATATGCTTCAAGCAATAATCTCAAC 780  
Db 880 TTCTATCTCTTCTGCAATAATCAAGCTTCAATACATCGCTCGGCAACAACTCAAC 939  
QY 781 ATCCCAATCTTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTTCATACGACGAGGCTC 840  
Db 940 ATCCCAATCTTTCAGTACCTTGATTCACAAAGCTTGGGGCTTCTTTCATACGACGAGGCTT 999  
QY 841 GATGAACACCGATGGAAGGAGATGTTCTCTATAGAGCTTGTCTCCATGGGCAATATA 900  
Db 1000 GACGAACTCCAGATGGAAGGAGCAATCTGTATCAGAGGCTTGTCTCCATGGGCAATATA 1059  
QY 901 GTTGAATTTACTTCGACAGCAATCTTGTGAGATCTTCTGGAAGTCTTCTGGAAGGCAACGTTCTAGG 960  
Db 1060 GTTGAATTTACTTCGACAGCAATCTTGTGAGATCTTCTGGAAGTCTTCTGGAAGGCAACGTTCTAGG 1119  
QY 961 AGTGAAGAAACCTCTTGTGCTCGGCGAGGACTTTTGTGAGTGTGTGATAGTACTCTGTCT 1020  
Db 1120 AGTGAAGAAACCTCTTGTGCTCGGCGAGGCTTCTGAGTGTGTGATAGTACTCTGTCT 1179  
QY 1021 ACCAATGTATCCCAAGACATCTTGTGAATATACTGTTGAATCTCTCTATGATCGCAATATC 1080  
Db 1180 TCCAAACCATCTCTGATCATCTGCTGATCTCTGCTGAGGCTCTCTGATGATCGGATATC 1239  
QY 1081 GAAGGTCTACTCAATGTTGAACACTGGGCAAACTTAAGAAATGAGAGCTGTGGAGT 1140  
Db 1240 GAAGGTCTACTCAATGTTGAACACTGGGCAAGCTTGGGCAAGAAATGAAAGTCTCTGGAGT 1299  
QY 1141 GTAGCAAGAGGTGTTATTAGAATGTTACGAAAAAACTATGTTGTTGTCGAGTGGATTTT 1200

Db 1300 GTGCAAGAGCGTATTACAGAAATGTCGGAAGAACTACCGCTATGTCCGAGTGGACTTT 1359  
QY 1201 GCACAGCATTTTCTTTAAAGGAATATTGAAAGACCAAGCTCAGAAACCGTGTCTGCT 1260  
Db 1360 GCACAGCATTTTCTTTGAGGAATATTTAGAGGCCAAAGTCAGAAACCTGTATCTGCT 1419  
QY 1261 CTACTTTCCCTGGAGCAAGCGTGTGTACCACTATATCTTCCATCAAGACCCAGTGTGCT 1320  
Db 1420 CCCCTCTCTTTGGAGCAAGCACTGTATACCACTAATCTTCTTCAAGACCTGTATGCTGCT 1479  
QY 1321 GCTGATGAGGTGAGACACGTCCTCAATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db 1480 GCTCCGAAACATAGAGACATGTCAGTAATAGTTCGAGAAACGCGGCGAGACGAAGCTTC 1539  
QY 1381 CGAAGGAGGTGTGTTGCAAAATCTGCGCTGAGCAATCTTATTCAGTCTAGCAAGTCTGCT 1440  
Db 1540 CGAAGGAGGTGTGTCGCAACCTGCGGAGACATCTTCTTCCCGCAGCAAGTCTCTGC 1599  
QY 1441 GCCATTATGTCCACACACATTTGTGGCTTGGCTGCTCTCTACAGACACAGGCAAGGAAT 1500  
Db 1600 GCTATCATGTCCACCCACATTTGTGGCTGCTGCTCTCTACAGACACAGGCAAGGAATC 1659  
QY 1501 GATCTCTCCATTTGGTGCAGACCTTCTTTGTGATGAAGAGGAAGTCTCGCTGCTGAT 1560  
Db 1660 CACCTCTCCACGCTGGTGAAGACTTCTTTGTGATGAAGAGGAAGTCTTACGTCTGGAT 1719  
QY 1561 TTTGACCTGGGTCTTCAGGAAATTCAGAAATGATAGTAAATGTCATGCCATACAGTCTGCTG 1620  
Db 1720 TTTGACCTGGGTCTTCTCGGGAATTCAGAAATGATAGTAAATGTCATGCTATTTTCACTGCTG 1779  
QY 1621 GGAATTTGTGTCACAATCACCCACACTAGCAGGAACATGATGATTTTATCACCCCGAGC 1680  
Db 1780 GGGAACTGTGTCACAATCACCCACACTAGCAGGAAGATGATTTTATTTACTCTCCAGC 1839  
QY 1681 ACAACTGTCCCATCAGTCTTGAAGTCTTCAAGCAATGAGGTAATCTCATGTCTTT 1740  
Db 1840 ACAACTGTCCGCTCGCTTGTGAAGTCTTCAAGCAATGAGGTAATCTCATGTCTTT 1899  
QY 1741 ATCATGAGGCCATCATAGTGTGAGCTTTATGAGTCTTGAACAGAGGGGACTGGGG 1800  
Db 1900 ATCATGAGGCCATCATAGTGTGAGCTTTATGAGTCTTGAAGCAATGAGGGGTTCCGGA 1959  
QY 1801 GGTCCCACTAGCACCCCACTTAACCTGATCAGCAGCAGCAGCTGTCGGAAGGCGGCC 1860  
Db 1960 GGGTCTCGGAGGCTTGGCAACTGATCAGCCAGAGCAGCTGTCGGAAGGCGGCC 2019  
QY 1861 AGCTGTGCTACTTCTCTCAATGAAGGCACTTCTCACTGCTGCTGCGACACATTTTAC 1920  
Db 2020 AGCTGTGCTACTTCTCTCAATGAAGTACCAATTTCTCTGCTGCTGCGACATTTTAC 2079  
QY 1921 CAAGTCTGCCATGAAACAGTAGGAAGTTTATCCAGTATGGCATTTCTACAGTGGCAGAG 1980  
Db 2080 CAGGTTTTGTCAAGAGACAGTAGGAAGTTTATCCAGTACCGAAATTTCTCACAGTGGCAGAG 2139  
QY 1981 CACCATCAGCAGGAAGATATCAGTCTGCTGCTGAGCAGCAGTGGCAAGAGCTT 2040  
Db 2140 CAAGTACCAAGGAAGATGTCAGTCTGCTGCTGAGCAGCAGTGGCAAGAGCTT 2199  
QY 2041 CCAGAACTTTGTCTTGGAGAAAGTATCAAGAAAGTGAAGACAGTGTGCTTTGGGAGGAA 2100  
Db 2200 CCGAGGCTCTGAACCTGGAGAGTGAAGAAAGTGAAGACAGTGTGCTTTGGTGGAGAG 2259  
QY 2101 CAGCGAATGCTTACCTGAAAGGTGAGCAATTCGAAGGAGCAGCAGAGTGTATCACCTTC 2160  
Db 2260 CAGCGTATGCTTACCTGAAAGGTGAGCAGGCCAAGGAGCAGCAGCAATTCATCACCTTT 2319  
QY 2161 TTACAGAGATCTCTTGGGCTTTGCTGAGGCTTACAGCTCTGCTGCACTTTGTTTAC 2220  
Db 2320 CTGACAGAGGCTTCTGGGGCTTCTGCTGAGGCTTACAGCTCTGCTGCACTTTGTTTAC 2379  
QY 2221 AACTTCAAGTGTCTCTGTTTCAGAACTGAGTATCTGCAAAAGTTCGCAAAATACCTAATA 2280

```
Db 2380 ACCTCCGGCCGACGTCCTCGAGTCTGAGTACTGCTGAGAAAGCTGCACAGGTACCTCTTC 2439
QY 2281 ACCAGAACAGAAAGAAATGTCAGTATATGCTGAGAGTGCCACATATGCTGTGTGAAG 2340
Db 2440 ACCAGAGGAGAGAAACGTCGGGTGTACCTGAGAGTGCCACATATGCTGTGTGAAG 2499
QY 2341 AATGCTGTGAAATGTTTAAAGSATAATGTTGGGTTTCAAGAGAGACCAAAACAAAGAGAGTG 2400
Db 2500 AATGCTGTGAAATGTTTAAAGGACATCGGGGTGTTTCAAGAGAGACCAAGCAGAGCGAGCG 2559
QY 2401 TCTGTTTGTAGAACTGAGCAGCAGCTTTCTACCTCAATGCAACCGACAAAACCTTCTAGAA 2460
Db 2560 TCTGTTTGTAGAACTGAGCAGCAGCTTTCTACCTCAAGGCGAGCGGAGAGCTCTCTGGAA 2619
QY 2461 TATATTCTGAGTTTGTGGTGCTG 2484
Db 2620 TACATTCTGAGCTTCGTGGTGCTG 2643

RESULT 12
BC019201
LOCUS
DEFINITION Mus musculus, glycerol-3-phosphate acyltransferase, mitochondrial,
Clone MGC:29133 IMAGE:5041120, mRNA, complete cds.
ACCESSION BC019201.1 GI:17512494
VERSION
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
Contact: MGC help desk
Email: gcaps@femail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nhgri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghigahi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 39 Row: k Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6680056.
Location/Qualifiers
1. 3857
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:29133 IMAGE:5041120"
/tissue_type="Kidney, normal. 5 month old male mouse."
/clone_lib="NCI_CGAP_Kid14"
/lab_host="DH10B"
```

## CDS

```
/note="Vector: pCMV-SPORT6"
160..2643
/codon_start=1
/product="glycerol-3-phosphate acyltransferase,
mitochondrial"
/protein_id="AAH19201.1"
/db_xref="GI:17512495"
/db_xref="LocustID:14732"
/translation="MEESGVTVGTIDVSLPSSSEYSLGRCKTSHEDWVDCGPKPTFF
RSATLKESLMRKRPFVGRCCYCTQSWERFPNPISPLGLRNVIVINETHTRHR
GHLARLSYLVQSRDVHKMFATSVENLVSSRVQBAIAEVAELNPDSQAQOS
KAIQKARKKILQEMVATVSPGMRILGTWLLKLFNSFVWNIQIHKQLQEMVKAT
ETNLLFLPVRHSIDYLLTFLPFCNFKAPYIAGNNFNIPIVSTIHLKGLGFI
RRLETDGPKDILYRALLGHVVELLRQQQLFLEGTSSRGKTCARAGLLSV
VVDLTSSNTIPDILVPIGVISYDRIIEGHYNGEQLGPKKNESLWSVARGVRLRKN
YGVVDFAPQPSLAKYLEGQSQVAPLSLEQALLPAILLPSRNDVADHDLLSSN
ESRNPADEAFRRRLIANLAEHLFTASKSCAIMSTHIVACLLLYRHRQIHSLVED
FFVKEEVLARDFDLFGNSDDVVMHAIQLLGNCTVITHTSRKDEFFITPSTVPSV
FELNFGVSLVHFIMEAIIACSIYVNLKRCSSGAGGLNLI SOEQLVKAASLCY
LLSNEGTTISLPQTFYQVCHETVGRFIOYQILTVAEQDDQEDVSQALQWMDKKLPE
LNWRDEDESDSGEORDCYLVSQSKHQOFTFLORLLGPLLEAVYSSAAIFVHN
FSGVPSEYLOKLHRYLITERNVAVYAESATYCLVKNVKNPKDIGVFKETQKR
VSVLSSTFLPQCHRKQLLEYLSEVVL"
BASE COUNT 941 a 952 c 1022 g 942 t
ORIGIN
Query Match 74.9%; Score 1860; DB 10; Length 3857;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 2101; Conservative 0; Mismatches 380; Indels 3; Gaps 1;
QY 1 ATGGATGAATCTGCACTGACCCCTTGGTCAATAGATGTTTCTTATCTGCCACATTCATCA 60
Db 160 ATGGAGGAGTCTTCAGTGACAGTGTGCACAAATAGAGCTTCTTATCTGCCAGTTCATCG 219
QY 61 GAATACAGTGTGGTGCATGTAAGCACACAACTGAGGAATGGGGTGCAGTGTGGCTTTAGA 120
Db 220 GAATACAGCTTGGCGGATGTAACACACACAGCAGTGGGTGGTGTGACTGTGGGTTCAA 279
QY 121 CCACCGTCTTCAGATCTCAACTTAAATGGAAGAAAGCTTAATAGTCGGAAGG 180
Db 280 CTTACCTCTTCAGATCTCAACACTGAATGGAAGAGAGCTTATGAGCGGAAGAGG 339
QY 181 CCATTTGTTGGAAGATGTTGTTTACTCTGCACTCCCGAGAGCTGGGACAAAATTTTCAAC 240
Db 340 CCCTTCGTGGGAAGGTGCTGCTATTCTCTGCACCGCACAGAGCTGGGAAAGGTTTTTCAAC 399
QY 241 CCAGTATCCCGTCTTGGGTTTGGGAATGTTATTTATATCAATGAAACTCACACAAG 300
Db 400 CCAGTATCCCATCTCTGGGTTTGGGAATGTTATTTATATCAATGAAACGCACACAAG 459
QY 301 CACCGGGATGCTTCAGAGAGCCCTTCTTACGTTCTTTTATTTTATTTTCAAGAGGAGATGTG 360
Db 460 CACAGAGGATGCTGGCGAGAGCGGCTGCTTACATCTCTTTTGTTCNAGAGCGAGAGCTC 519
QY 361 CATAAGGGCATGTTTGGCCACCAATGCTGCTGAAAATGCTGCTGAACAGCAGTAGAGTACAA 420
Db 520 CATAGGGCATGTTTGGCCACCATGTTACTGAGATGTTACTGAGCAGCAGCAGAGTCCAA 579
QY 421 GAGGCAATTCAGAGAGTGGCTGCTGAATTAACCCGTGATGGTGTCTGCCCGCAGAGCAATCA 480
Db 580 GAGGCAATTCGTGAAGTGGCTGGAGTTGAAACCCAGATGGATCTGCCCGCAGCAGAGTCC 639
QY 481 AAAGCGTTTAAACAAGTGAAGAAAGCTAAAGAGTTCTTCAAGAAATGTTGTCACCT 540
Db 640 AAAGCGCATCCAGAGGTGAAGAAAGGAAAGCCAGGAAGATCCTCCAGAGATGTTGTCACCC 699
QY 541 GTCTCACCCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTGCTTAAACTGTTCAACAGCTTC 600
Db 700 GTCTCCCGCAGGATGATCAGGCTGACTGGCTGGGTGTTTAAAGCTCTTCAACAGCTTC 759
QY 601 TTTTGAACATTTCAAAATTCACAAAGGTCAACTTGAGATGGTTAAAGCTGCACTGAGAGC 660
Db 760 TTTTGAACATTTCAATTCACAAAGGTCAACTCGAGAGGTGTCAGGCTGCACTGAGAGC 819
```

QY 661 AATTGCGGCTTCTGTTTCTACCACTTCATAGATCCCATATTTGACTATCTGCTGCTACT 720  
Db |||||  
QY 820 AACCTGCGGCTCTGTTTCTGCGGTGCACAGATCCCAATTTGACTACTCTGTTGCTCACC 879  
Db |||||  
QY 721 TTCAATCTCTCTGTCATTAACATCAAGACCAATACATTTGCTTCAAGCAATAATCTCAAC 780  
Db |||||  
QY 880 TTCACTCTCTTTTGGCCACAACATCAAGGCGCGTACATCGCTCGGGCAATAATCTCAAC 939  
Db |||||  
QY 781 ATCCCAATCTTCAGTACCTTTCATCATTAAGCTTTGGGGCTTCTTCATACGACGAAGGCTC 840  
Db |||||  
QY 940 ATCCCGCTCTTCAGTACCTTTCATCAAGCTTTGGGGCTTCTTCATAGACGAGGCTC 999  
Db |||||  
QY 841 GATGAAACACCAAGATGAGCGGAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db |||||  
QY 1000 GATGAAACCCAGATGAGCGGAAAGACATCTGTACAGAGCTTCTCCATGGGCATGTA 1059  
Db |||||  
QY 901 GTTGAAATTTACGACAGCAGCAATTTCTTGAGATCTTCCTGGAAGGACACAGTCTTAGG 960  
Db |||||  
QY 1060 GTTGAAATCTCCGACAGCAGCAGTTCCTGGAAGGACACCGCTCCCGC 1119  
Db |||||  
QY 961 AGTGAAAAACCTCTGTGCTCGGCGAGGACTTTTGTCAAGTTGTGTAGATCTCTGCT 1020  
Db |||||  
QY 1120 AGTGCAAGACTCTGTGCGCGGCGAGGCTCTCTCAGTGTAGTGTGATCTCTGTGCG 1179  
Db |||||  
QY 1021 ACCAATGTCTCCAGACATCTTGATATAATCCCTGTGGAATCTCTATGATCGCATATC 1080  
Db |||||  
QY 1180 TCCAACCATCTCCGACATCTCTGTCATACCCGTGGGCATCTCTGATGTCGATATC 1239  
Db |||||  
QY 1081 GAAGTCTACTACATGGTGAACAACTGGGCAAACTTAAGAAATGAGAGCTGTGGAGT 1140  
Db |||||  
QY 1240 GAAGTCTACTCAATGGGCAACAGTTGGGAAGGCCAAGAAAGACGAGAGCTCTGGAGT 1299  
Db |||||  
QY 1141 GTAGCAAGAGGTGTTATTAGAATGTTACGAAAACTATGTTGTGTCGAGTGGATTTT 1200  
Db |||||  
QY 1300 GTGGGAGAGGGTTATCAGAAATGCTGCGGAAAACTACGGCTACGTCGAGTGGATTTT 1359  
Db |||||  
QY 1201 GCACAGCCATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGTCGCT 1260  
Db |||||  
QY 1360 GCACAGCCATTTTCTTGAAGGAATATTTAGAGGCGAGATCGAAGACCTGTATCTGCC 1419  
Db |||||  
QY 1261 CTACTTTCCCTGGAGCAAGCTTGTTCAGAGTATATCTTCTTCAAGACCCAGTGTGCT 1320  
Db |||||  
QY 1420 CCCCTTTCTCTGGAGCAAGCACTGTTCACGCGATCTTCTTCAAGACCCGAAATGATGT 1479  
Db |||||  
QY 1321 GCTGATGAAGGTAGAGACAGTCCATTAATGAGTCCAGAAATGCAACAGATGAAATCCCTA 1380  
Db |||||  
QY 1480 GCTGATGAACATCAAGACCTATCCAGTAACGAGTCCAGAAACCCAGCAGACGAAGCCTTC 1539  
Db |||||  
QY 1381 CGAAGGAGGTGATTCGAAATCTGGCTGAGCATATTTCTATTCTGCTAGCAAGTCTCTGT 1440  
Db |||||  
QY 1540 CGACGAAGGCTGATTCGAAACCTGGCTGAGCACAATTTCTTTCACCGCCAGCAAGTCTGCG 1599  
Db |||||  
QY 1441 GCCATATGTCCACACACATTTGGCTGCTGCTCTCTCTACAGACACAGGAGGGAAT 1500  
Db |||||  
QY 1600 GCTATCATGTCCACCACATTTGGGCTGTCTGCTCTCTCTACAGACACAGGAGGGAATC 1659  
Db |||||  
QY 1501 GATCTCTCCACATTTGTCGAGACTTCTTGTGATGAAGGAGGAGTCTGCTGCTGTAT 1560  
Db |||||  
QY 1660 CATCTCTCCAGCTTGTGGAAGACTTCTTGTGATGAAGGAGGAGTCTTACGCTCGCAT 1719  
Db |||||  
QY 1561 TTTCAGCTGGGGTCTTCAGGAAATTCAGAAAGATGTAGTAATGCATGCGCATACAGCTGCTG 1620  
Db |||||  
QY 1720 TTGACCTAGGCTTCTCCGGGAATTCAGAAAGTGTGTCATGTCATGCTATTTCAGCTTCTG 1779  
Db |||||  
QY 1621 GGAATTTGTTCACAAATCACCCACACTAGCGAAACGATGATGTTTTTATCACCCCGCAGC 1680  
Db |||||  
QY 1780 GGGAACTGTGTACAAATCACCCACAGCAGGAGGAAAGATGATTTTTTATTACTCCACG 1839  
Db |||||  
QY 1681 ACAACTGTCCCATGATCTTGAACCTCACTTACAGCAATGGGGTACTTCAATGCTTTT 1740  
Db |||||  
QY 1840 ACAACTGTCCGCTAGTCTTTGAACCTCACTTACAGCAATGGGGTACTTCAATGCTTTC 1899

QY 1741 ATCATGGAGGCCATCATAGCTTGCAGCTTTATGCACTTCTTGAACAAGAGGGGACTGGGG 1800  
Db |||||  
QY 1900 ATCATGGAAGCCATCATAGCTTGCAGCATCTATGCACTCTCTGAATAAGAGGTGCTCTGGA 1959  
Db |||||  
QY 1801 GGTCCCACTAGCACCCACCTAACTGATCAGCAGGAGGAGCTGGTCCGGAAGCGGCC 1860  
Db |||||  
QY 1960 GGGTCCGCTGGAGCCCTCGGCAACTGATCAGCAGGAGGAGCTGGTGGGAAGCGCC 2019  
Db |||||  
QY 1861 AGCTGTGCTACCTTCTCTCAATGAAGGACCACTCTCACTGCTTGGCCATGCGACATTTTAC 1920  
Db |||||  
QY 2020 AGCTGTGCTACCTTCTCTTAAGGAAGTACATTTCTGCTGCTGCGCTGCGACATTTTAC 2079  
Db |||||  
QY 1921 CAAGTCTGCCATGAAACAGTAGGAAGTTTATCAGTAGTGCATTTCTTACAGTGGCAGAG 1980  
Db |||||  
QY 2080 CAAGTTTGTCTATGAGACAGTTGGCAAGTTTATCAGTATGCAATTTCTCACAGTGGCAGAG 2139  
Db |||||  
QY 1981 CACGATGACCAAGGAAGATATCAGTCTCTAGTCTTCTGAGGAGCAGTGGGACAAAGAGCTT 2040  
Db |||||  
QY 2140 CAAGATGACCAAGGAAGATGTCAGTCTCTGCTTCTGAGGAGCAGTGGGACAAAGAGCTT 2199  
Db |||||  
QY 2041 CCAGAACCTTTGCTTGGAGAAAGTATGAAGAGATGAAGACAGTGCATTTTGGGAGGAA 2100  
Db |||||  
QY 2200 CCGAAGC---TGAAGTGGAGAAAGTGAAGGAAGATGAAGACAGTGCATTTTGGTGGAGAG 2256  
Db |||||  
QY 2101 CAGCAGATGCTACTACCTGAGGCTGAGCCAAATCCAAAGGAGCACAGCAGTATTTATCACCTTC 2160  
Db |||||  
QY 2257 CAGCAGATGCTATCTCAAGTGCAGCAGTCCAAAGGAGCACAGCAATTCATCACCTTT 2316  
Db |||||  
QY 2161 TTACAGAGACTCTTGGCCCTTCTGAGGAGCCTTACAGTCTCTGCTGCTGCTGCTTTCAC 2220  
Db |||||  
QY 2317 CTTACAGAGGCTTCTAGGTCCTTCTGTAAGACCTTACAGTCTGCTGCTGCTGCTTTCAC 2376  
Db |||||  
QY 2221 AACTTTCAGTGTCTGTTTCCAGAACCTGAGTATCTGCAAAAGTTCGCAAAATACCTAATA 2280  
Db |||||  
QY 2377 AACTTTCAGCGTGCAGTTCCTGAGTCTGCAAAAGTTCGCAAAAGTTCGCAAAAGTTCATC 2436  
Db |||||  
QY 2281 ACCAAGACAGAAAGAAATGTTGCAAGTATATCTGAGAGTGCACATATTTCTTGTGAAG 2340  
Db |||||  
QY 2437 ACCAGGACGGAAGAAAGACGTTGCGGTGAGTCTGAGAGTGCACATATCTGCTGCTGAGAG 2496  
Db |||||  
QY 2341 AATGCTGTGAAATGTTTAAAGGATATTTGGGTTTCTAAGGAGACCAACAAAGAGAGTG 2400  
Db |||||  
QY 2497 AACGCTGTGAAATGTTTAAAGGACATCGGGTCTTCAAGAGAGACCAAGCAAAAGCGAGTG 2556  
Db |||||  
QY 2401 TCTGTTTTCAGAACGAGCAGCAGTCTTCTACTCAATGCAACCCGACAAATCTTCTAGAA 2460  
Db |||||  
QY 2557 TCTGTTTCAGAACGAGCAGTCTTCTACTCAATGCAACCCGACAAATCTTCTAGAG 2616  
Db |||||  
QY 2461 TATATTCAGTGTGTTGTTGCTGCTG 2484  
Db |||||  
QY 2617 TATATTCAGTGTGTTGTTGCTGCTG 2640  
Db |||||

RESULT 13  
MUSG3PAT

LOCUS 6634 bp mRNA linear ROD 27-APR-1993  
DEFINITION Mouse glycerol-3-phosphate acyltransferase mRNA, complete cds.  
ACCESSION M77003  
VERSION M77003.1 GI:193366  
KEYWORDS glycerol-3-phosphate acyltransferase.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 6634)  
AUTHORS Shin,D.H., Paulauskis,J.D., Moustaid,N. and Sul,H.S.  
TITLE Transcriptional Regulation of p90 with sequence homology to E. coli  
J. Biol. Chem. (1991) In press  
JOURNAL 2 (sites)  
AUTHORS Shin,D.H., Paulauskis,J.D., Moustaid,N. and Sul,H.S.  
TITLE Transcriptional regulation of p90 with sequence homology to  
Escherichia coli glycerol-3-phosphate acyltransferase









BASE COUNT		proteases"		627 a		670 c		644 g		578 t		Db	QY
ORIGIN													
Query Match		72.6%;		Score 1803;		DB 10;		Length 2519;					
Best Local Similarity		84.4%;		Pred. No. 0;									
Matches 2041;		Conservative		0;		Mismatches 375;		Indels 3;		Gaps 1;			
QY	66	CAGTGTGGTTCGATGTAAGCACACACAGTGAAGGATGGGTGAGTGTGGCTTTAGACCCAC	125									961	CACCATCCCTGACATCCCTGTGGGCATCTCGTATGATCGGATAATCGAAGG
Db	1	CAGCCTTGGCCGATGTAAACACACAGATGAGGACTGGGTTCGATTTGGCTTTCAACCTTAC	60									1086	TCACTACAAATGGTGAACAACTGGGCAAACTTAAGAAGAAATGAGAGCCTCTGGAGTGTAGC
QY	126	CGTCTTCAGATCTGCACCTTTAAATGGAAGAAAGCCTTAATGAGTCGGAAAGGCGCATTT	185									1021	TCACTACAAATGGTGAACAACTGGGCAAGCCCAAGAGNAATGAAAGTCTCTGGAGTGTGGC
Db	61	CTTCTTCAGATCCGCAACGCTGAAATGGAAGGAGAGCCTCATGAGCCGGAAGAGGCGCCTT	120									1146	AAGAGGTGTATTAGAATGTTTACGAAATAAATACTATGTTGTGTCGAGTGGATTTTGACAC
QY	186	TGTTGGAGATGTTGTTACTCTGCACTCCCGAGAGCTGGGACAAATTTTTCACCCCGAG	245									1081	AAGAGGCGTTATCAGAAATGCTCGGCAAAAACCTACGGCTATGTCGCGAGTGGACTTTTGACA
Db	121	CGTGGGAAGGTGTGCTATTATGCAACGCTCAGAGCTGGGAAGGTTCCTTCAACCCCGAG	180									1206	GCATTTTCTTAAAGGAATATTTAGAAAGCAAAGTCAAGAACCGGTGTCTGCTCTACT
QY	246	TATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAATACTCACACAGACACCG	305									1141	GCATTTTCTTGAAGGAATATTTAGAAAGGCCAAAGTCAAGAACCTGTATCTGTCTCCCT
Db	181	TATCCCACTCTGGGTTTGGGNAATGTTATTTATATCAATGAATACTCACACAGGACCG	240									1266	TTCCCTGGAGCAAGCGTTGTACAGACTATATCTTCTTCAAGACCCAGTGTGCTGCTGA
QY	306	CGATGGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGAGCGAGATGTGCATAA	365									1201	CTCTTTGGAGCAAGCACTGTGTACCAAGCATCTCTTCTTCAAGACTGATGCTGCTGCTGC
Db	241	AGGATGGCTGGCAAGAGCGCTTCTTACATCTCTTTTGTTCAGAGCGGATGTCCACAA	300									1326	TGAAGGTGAGACACGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAG
QY	366	GGCAGTGTGGCACCAATGTGACTGAAATGTGCTGAAACAGAGTAGTATGATACAGAGGC	425									1261	CGAACATGAAGACATGTCCAGTAAATGAGTCGAGAAACGCGGCAGACGAAAGCTTCCGAAG
Db	301	GGGATGTTTGGCACCAATGTGACTGAAATGTGCTGAAACAGAGTAGTATGATACAGAGGC	360									1386	GAGGTGATTTGCAAAATCTGGCTGAGCATATTTCTATTTCACTGTAGCAAGTCTCTGTGCCAT
QY	426	AATTGAGAGTGGCTGCTCAATTAACCTGTATGTTCTGCCCAGCAGCAATCAAAAGC	485									1321	GAGGCTGATCGCAAACTCGCGGAGCAAACTCTCTCACCGCCAGCAAGTCTCTGCGCTAT
Db	361	AATTGCTGAAGTGGCTGCAAGATTGAAACCGGATGATCTGCCAGCAGCAGTCCAAAGC	420									1446	TATGTCCACACATTTGTGGCTTGGCTCTCTCTACAGACACAGGCAAGGAATTTGATCT
QY	486	CGTTAAACAAAGTGAAGAAAGCTTAAAGGATTTCTTCAAGAAATGGTTGCCACTGTCTC	545									1381	CATGTCCACCCCACTTGTGGCTTGGCTCTCTCTCTACAGACACTGGCAGGGAATCCACCT
Db	421	CATCCAGAAAGTGAAGAAAGCCAGGAAGATCTCCAGGAAATGGTTGCTACAGTCTC	480									1506	CTCCACATTTGGTCGAAGACTTCTTTGTGATGAAAGAGGAAGTCTCTGGCTCGTGAATTTGA
QY	546	ACGGCAATGATCAGACTGACTGGGTGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	605									1441	CTCCAGCTTGTGGAAGACTTCTTTGTGATGAAAGGGAAGTCTCTAGCTCGGGAATTTGA
Db	481	CCCGGGATGATCAGGCTGACTGGCTGGGTGTTACTTAAAGCTCTTTCACAGCTTCTTCTG	540									1566	CCTGGGTTTCTCAGGAAATTCAGAAAGTGTAGTAAATGCAATGCCATACAGCTGCTGGGAAA
QY	606	GAACATTCAAATTCACAAAGGTCACTTCAGATGTTTAAAGCTGCAACTGAGAGCGAATTT	665									1501	CCTGGGCTTCTCCGGAATTCAGAAAGCCTGATGATGATGATGATGATGATGATGATGATGAT
Db	541	GAACATTCAGATTCACAAAGGTCACTTGAGATGTTGAAAGCTGCACACTGAGAGCAATCT	600									1626	TTGTGTCAAAATCAACCCACACTAGCAGGAAACGATGATGATGATGATGATGATGATGATGAT
QY	666	GCCTCTTCTGTTTCTACAGTTTCATAGATCCCATATTTGATCTATCTGCTGCTGCTGCTGCTG	725									1561	CTGTGTCAAAATCAACCCACACTAGCAGGAGAGTAAATCTTTTATTTACTCCAGACAAAC
Db	601	GCCTCTTCTGTTTCTGCGGTGCACAGATCCCATACGACTACCTGTTGCTCACCTTCAT	660									1686	TGTCCCATCAGTCTTCGAACCTCACTTCTACAGCAATGGGGTACTTCTATGCTTTTATCAT
QY	726	TCCTCTTCGCATAACATCAAAAGCACCATACATGCTTTCAGGCAATTAATCTCAACATCCC	785									1621	TGTCCCGTCCGTTTGTGAACCTCAACTTCTACAGCAATGGGGTACTTCTATGCTTTTATCAT
Db	661	CCTCTTCTGCTCAACATCAAAAGCTCCATACATCGCTCGGCAACCACTCAACATCCC	720									1746	GGAGGCCATCATAGCTTGCAGCTTTTATGCACTTCTGAAACAAAGAGGGAACTGGGGGTCC
QY	786	AATCTTCAGTACCTTGCATCAAGCTTGGGGCTTCTTCTATACGACGAAGGCTCGATGA	845									1681	GGAGGCCATCATAGCTTGCAGCTTTTATGCACTTCCAGATTAAGAGGGTTCGCGAGGTCT
Db	721	CATCTTCAGTACCTTGAATTCACAGCTTGGGGCTTTTTCATAAGCAGGAGCTTGAGCA	780									1806	CATAGCACCCACCTTAACCTGATCAGCAGGAGCAGCTGGTCCGGAAGCGCGCCAGCCT
QY	846	AACACAGATGGAGGAAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCAATATATAGTTGA	905									1741	GCC---GGAGCCCTGGCAACCTGATCAGCAGGAGAGCTGGTCCGGAAGCGCCAGCCT
Db	781	AACTCCAGATGGAGCNAAGACATCTGTACAGAGGTTGCTCCATGGGCAATGATTTGA	840									1866	GTGCTACCTTCTCTCAATGAAGGACCACTCTCACTGCTTGCCTTGCAGACATTTTACCAAGT
QY	906	ATTACTTCGACAGCAATCTTGGAGATCTTCTTGGAGGCAACGTTCTAGAGTGG	965									1798	GTGCTACCTTCTCTTAATGAAGTACCAATTTCTCTGCCCTGCCAGACATTTTACAGGT
Db	841	ACTCTCTCCGACAGCAGTCTCTGAGATCTTCTCTGGAAGGCAACCGCTCCCGCAGTGG	900									1926	CTGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCAATTTTACAGTGGCAGAGACGA
QY	966	AAAACTCTTGTCTGGGCGAGCTTTTGTGATGTTGTTGATGATGATGATGATGATGATGATG	1025									1858	TTGTCCAGAGACAGTAGGAAAGTTTATCCAGTACCGAAATCTCACAGTGGCAGAGCAAGA
Db	901	CAAGACCTCTGTGCTGGGCGGCGCTCTGTGATGATGATGATGATGATGATGATGATGATG	960									1986	TGACCGAGGAGATATCAGTCTTCTGCTGAGCAGAGTGGGCAAGAGCTTCCAGA
QY	1026	TGTCATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATGCGCATTAATCGAAG	1085									1918	TGACCGAGGAGATGTCAGTCTCTGGCCTTTGCAGAGCAGCAGGTGGAAACAAAGACTTCCGGA
Db	1026	TGTCATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATGCGCATTAATCGAAG	1085									2046	ACCTTTGTCTTGGAGAGTGAAGAAAGTGAAGACAGTGAATTTTGGGGAGGAACAGCG
QY	1026	TGTCATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATGCGCATTAATCGAAG	1085									1978	GCCTCTGAACCTGGAGAGTGAAGAGAGATGAGGACAGTGAATTTGGTGGGAGAGAGCG
Db	1026	TGTCATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATGCGCATTAATCGAAG	1085									2106	AGATTGCTACCTGAAGGTGAGCCAAATCCAAAGGACCAACGAGTTTATCACCTTCTTTACA

Db	2038	TGATTGCTACCTGAAGGTGAGCCAGGCGAAGGAGCACAAGAGAAATTCATCACTTTCTTGCA	2097
Qy	2166	GAGACTCCTTTGGGGCTTTTGCTGGAGGCGCTACAGCTCTGCTGCCATCTTTGTTCAACAATTT	2225
Db	2098	GAGGCTTCTGGGGCCCTGCTGTAGAAGCCTACAGCTCTGCTGCCATCTTTGTTCAACACTTT	2157
Qy	2226	CAGTGTCTCTGTTTCCAGAAACCTGAGTATCTGCAAAAGTTGCGACAAATACCTAATAACCAG	2285
Db	2158	CCGCGCGCCAGTTCCTCCGAGTCTGAGTACCTGCGAGAGCTGCGACAGGTACCTTCTCACCG	2217
Qy	2286	AACAGAAGAAATGTTCCAGCTATATGCTGAGAGTGCCACATATTTGTTCTGTGAAGATGC	2345
Db	2218	GACGGAGGAGACGTCGCGGTGTACGTCGAGAGTGCCACATCTGTCTTTGTGAAGAATGC	2277
Qy	2346	TGTCAAAATGTTTAAAGGATATTGGGGTTTTCAAGGAGACCAAAACAAAAGAGAGTGTCTGT	2405
Db	2278	TGTGAATAATGTTTAAAGACATCGGGGTTTTCAAGAGACCAAGCAGGAAGCGAGCGTCTGT	2337
Qy	2406	TTTAGAACTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAAACTTCTAGAAATATAT	2465
Db	2338	CTTAGAACTGAGCACCACTTTCCTACCTCAGGCGACCGGAGAGCTCTCGAATACAT	2397
Qy	2466	TCTGAGTTTTGTGTGCTG	2484
Db	2398	TCTGAGCTTGTGTGCTG	2416

RESULT 15				
AL391986/c				
LOCUS	AL391986	196657 bp	DNA	linear
DEFINITION	Human DNA sequence from clone Rp11-426E5 on chromosome 10, complete sequence.			

ACCESSION AL391986  
VERSION AL391986.12 GI:11544543

KEYWORDS	HTG.
SOURCE	Homo sapiens (human)

ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 196657)

**AUTHORS** Ashwell, R.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (30-NOV-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Dec 4, 2000 this sequence version replaced gi:11414640. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr10>

RP11-426B5 is from the library RPC1-11.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/> VSCITOR: PBACe3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-426B5 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
1	53	2.1	2598	4	US-09-252-991A-4323	Sequence 4333, Appl	
C	53	2.1	2874	4	US-09-252-991A-4533	Sequence 4533, Appl	
	46.2	1.9	2589	4	US-09-328-352-2160	Sequence 2160, Appl	
C	4	44.8	1.8	1830121	4	US-09-557-884-1	Sequence 1, Appl
C	5	44.8	1.8	1830121	4	US-09-643-990A-1	Sequence 1, Appl
C	6	38.6	1.6	1664976	4	US-08-916-421B-1	Sequence 1, Appl
C	7	37.2	1.5	1578	4	US-09-328-352-2557	Sequence 2557, Appl
C	8	36.6	1.5	2790	3	US-08-895-601-2	Sequence 2, Appl
C	9	35.6	1.4	7218	1	US-08-323-463-14	Sequence 14, Appl
C	10	35.2	1.4	1032	4	US-09-328-352-2440	Sequence 2440, Appl
C	11	34.8	1.4	202001	4	US-09-734-674-3	Sequence 3, Appl
C	12	34.2	1.4	825	4	US-09-328-352-3343	Sequence 3343, Appl
C	13	34	1.4	1711	1	US-08-749-903-2	Sequence 2, Appl
C	14	34	1.4	1711	4	US-09-088-641-2	Sequence 2, Appl
C	15	34	1.4	9046	1	US-08-327-536-1	Sequence 1, Appl
C	16	34	1.4	9046	5	PCT-US95-0468B-1	Sequence 1, Appl
C	17	33.6	1.4	2774	3	US-08-466-548B-1	Sequence 1, Appl
C	18	33.6	1.4	2774	3	US-07-998-226F-1	Sequence 1, Appl
C	19	33.6	1.4	2774	5	PCT-US93-12560-1	Sequence 1, Appl
C	20	33.6	1.4	4821	3	US-08-913-374-1	Sequence 1, Appl
C	21	33.4	1.3	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C	22	33.4	1.3	4411529	3	US-09-103-840A-1	Sequence 1, Appl
C	23	33.2	1.3	1534	4	US-08-858-207A-93	Sequence 93, Appl
C	24	33.2	1.3	1848	4	US-09-134-001C-447	Sequence 447, Appl
C	25	33.2	1.3	7672	4	US-09-220-132-24	Sequence 24, Appl
C	26	33	1.3	3752	4	US-08-961-527-208	Sequence 208, Appl
C	27	33	1.3	319608	4	US-09-539-333D-1	Sequence 1, Appl





```
;/ NAME/KEY: misc feature
;/ LOCATION: (103998)..(103998)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (148948)..(148948)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (163385)..(163385)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (191989)..(191989)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (191995)..(191995)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (231980)..(231980)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (234187)..(234187)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (234220)..(234220)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (234814)..(234814)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (309398)..(309398)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (309418)..(309418)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (312837)..(312837)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (312993)..(312993)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (319226)..(319226)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (559167)..(559167)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (559241)..(559241)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (600992)..(600992)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (622708)..(622708)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (657081)..(657081)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (657203)..(657203)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (674435)..(674435)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (682442)..(682442)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (713652)..(713652)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (741684)..(741684)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
```

```
;/ LOCATION: (779455)..(779455)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (779676)..(779676)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (855539)..(855539)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (871619)..(871619)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1084830)..(1084830)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1096846)..(1096846)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1119881)..(1119881)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1130881)..(1130881)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1310988)..(1310988)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1313224)..(1313224)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1349473)..(1349473)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1349491)..(1349491)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1470091)..(1470091)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1569020)..(1569020)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1602912)..(1602912)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1603734)..(1603734)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1637998)..(1637998)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ NAME/KEY: misc feature
;/ LOCATION: (1664854)..(1664854)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

Query Match 1.6%; Score 38.6; DB 4; Length 1664976;

Best Local Similarity 50.8%; Pred. NO.14;

Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```
QY 338 TTTTATTCAAGAGCGAGATGTGCATAAGGGCATGTTGCCCAATGTGACTGCTGAAATG 397
Db 541143 TCTTTATTAAACATAGAGAGGATAACAGATGCATGATATAGTAGTGAATGTCATG 541084

QY 398 TCGTGAACAGCAGTAGAGTACAAGAGGCAATTCGCAAGTGGCTGCTGAATTAACCCCTG 457
Db 541083 CTCTTAAATGAAATATTAAAGAAATAACAAAATTTCTTCAGTTGTAGGAGAGCCAA 541024

QY 458 ATGGTTCTGCCGAGCGCAATCAAAAGCGTTAAACAAAGTGAAGAAAGAAAGCTAAAGGA 517
Db 541023 AAACCTCTGGAAGTTGGATACAATAATGTTGGTGAATGTAAGAGAGCATGTGAAAACTA 540964

QY 518 T 518
```





```
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ptzgt-Fls
US-08-232-463-14

Query Match      1.4%; Score 35.6; DB 1; Length 7218;
Best Local Similarity 1.5%; Pred. No. 2.8;
Matches 5; Conservative 185; Mismatches 134; Indels 0; Gaps 0;

QY 339 TTTTATTCAGAGCGAGATGTCATAAGGCGCATGTTTGGCCACCAATGTCGCTGAAAATGT 398
Db 1442 TTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1383
QY 399 GCTGAACAGCAGTAGTCAAGAGGCAATTCGAGAAAGTGGCTGCTGCAATTAACCCCTGA 458
Db 1382 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1323
QY 459 TGGTCTGCCACAGACAAATCAAAAGCGTTAAACAAAGTGAAAGAAAGAGCTAAAGGAT 518
Db 1322 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1263
QY 519 TCTTCAAGAAATGGTCCCACTGCTCACCGCAATGATCAGACTGACTGGGTGGTGCT 578
Db 1262 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1203
QY 579 GCTAAAACTGTTCAACAGCTTCTTTTGGAACTTCAAAATTCACAAAGGTCAACTTGAGAT 638
Db 1202 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1143
QY 639 GGTAAAGCTGCAACTGAGACGAA 662
Db 1142 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1119

RESULT 10
US-09-328-352-2440/c
; Sequence 2440, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2440
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2440

Query Match      1.4%; Score 35.2; DB 4; Length 1032;
Best Local Similarity 48.5%; Pred. No. 0.97;
Matches 97; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 342 TATTACAGCGAGATGTCATAAGGCGCATGTTTGGCCACCAATGTCGCTGAAAATGCT 401
Db 425 TATTCCAGCTCTTTAAATTTGATCATTTGAATCAATTAATGCAATTTTGAATGGTGTCT 366
QY 402 GAACAGCAGTAGAGTACAAGAGGCAATTCGAGAAAGTGGCTGCTGAAATTAACCCCTGATGG 461
Db 365 TGACCAATCAGCGATGTAACACCGCTGTTTAAATCGGATCTTGAATAAATTCACGGTGG 306
QY 462 TTCTGCCACAGCAATCAAAAGCGGTTAAACAAAGTGAAAGAAAGAGCTAAAGGATTTCT 521
Db 305 TGACGAGCTGTTTCTATCACCAGATCTTCACCATAGTCAAAACCAACCGAAGCGTTGCA 246
QY 522 TCAGAAATGGTGGCACTG 541
Db 245 AAGAACAAAGTTGCAAACTG 226
```

```
RESULT 11
US-09-734-674-3
; Sequence 3, Application US/09734674
; Patent No. 6498022
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001018
; CURRENT APPLICATION NUMBER: US/09/734,674
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 202001
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(202001)
; OTHER INFORMATION: n = A,T,C or G
US-09-734-674-3

Query Match      1.4%; Score 34.8; DB 4; Length 202001;
Best Local Similarity 60.6%; Pred. No. 54;
Matches 57; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 2216 TTCACAACTTCAGTGGTCTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAATACC 2275
Db 32955 TTCTCTAGTCTTTAGGCCAGTGTTATATGTTAGGATTTACAAAAGTTGGTAATATAGA 33014
QY 2276 TAATACCAAGACAGAGAAATGTTGCAGTATA 2309
Db 33015 GAGAAACAGGAGAAATGAAATGGACAGGAAA 33048

RESULT 12
US-09-328-352-3343/c
; Sequence 3343, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3343
; LENGTH: 825
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3343

Query Match      1.4%; Score 34.2; DB 4; Length 825;
Best Local Similarity 52.4%; Pred. No. 1.7;
Matches 75; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 1348 AATGATGTCAGAAAATGCAACAGATGAATCCCTACGAAGGAGGTGATTGCAAAATCTGGCT 1407
Db 364 ATTGATGTAACAAGGCGAGACTTGTATAACCTCGCAATAGTTTGTCTCTTTGGGCGAT 305
QY 1408 GAGCATATTCTATTCTACTGCTAGCAGTCTGTGCGCATTTATGTCACACACATTTGGCT 1467
Db 304 TACTCATTCTTATATACAACTATAAGTACTGACTTTAGATAACCATACCTGATGTGGTT 245
QY 1468 TGCCTGCTCTCTACAGACACAG 1490
Db 244 TGCCTGCCGGTATCCATACCG 222

RESULT 13
```

US-08-749-903-2

Sequence 2, Application US/08749903  
Patent No. 5759812  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Hawkins, Phillip R.  
TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ Version 1.5  
CURRENT APPLICATION NUMBER: US/09/088,641  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/749,903  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0163 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6312895e  
IMMEDIATE SOURCE:  
CLONE: 989953

US-09-088-641-2

Query Match 1.4%; Score 34; DB 1; Length 1711;  
Best Local Similarity 63.4%; Pred. No. 3.4;  
Matches 52; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1258 GCTCTACTTTCCTGGAGCAGCGTTGTACACGTATATCTTCAAGACCCAGTGAT 1317  
DB 1270 GCTGTACAGTGCCTGGGCAAGCAGTTTACCCCTGATCTCATCAGGAAGGCTCTGTGAT 1329

QY 1318 GCTGCTGATGAAGGTAGACACA 1339  
DB 1330 GCTGCAGGTTGATGTAGACACA 1351

RESULT 15

US-08-227-536-1/c  
Sequence 1, Application US/08227536  
Patent No. 5658784  
GENERAL INFORMATION:  
APPLICANT: Eckner, Richard  
APPLICANT: Ewen, Mark  
APPLICANT: Livingston, David  
TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION  
TITLE OF INVENTION: FACTOR P300 AND USES OF P300  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
STREET: Ten Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/08/227,536  
FILING DATE: 14-APR-1994  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Ph.D., Kathleen A.  
REGISTRATION NUMBER: 34,380  
REFERENCE/DOCKET NUMBER: DFCI-308XX

US-08-749-903-2

Query Match 1.4%; Score 34; DB 1; Length 1711;  
Best Local Similarity 63.4%; Pred. No. 3.4;  
Matches 52; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1258 GCTCTACTTTCCTGGAGCAGCGTTGTACACGTATATCTTCAAGACCCAGTGAT 1317  
DB 1270 GCTGTACAGTGCCTGGGCAAGCAGTTTACCCCTGATCTCATCAGGAAGGCTCTGTGAT 1329

QY 1318 GCTGCTGATGAAGGTAGACACA 1339  
DB 1330 GCTGCAGGTTGATGTAGACACA 1351

RESULT 14

US-09-088-641-2

Sequence 2, Application US/09088641  
Patent No. 6312895  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Hawkins, Phillip R.  
TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ Version 1.5  
CURRENT APPLICATION NUMBER: US/08/749,903  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0163 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 5759812e  
IMMEDIATE SOURCE:  
CLONE: 989953

US-08-749-903-2

US-08-749-903-2

Sequence 2, Application US/08749903  
Patent No. 5759812  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Hawkins, Phillip R.  
TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ Version 1.5  
CURRENT APPLICATION NUMBER: US/09/088,641  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/749,903  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0163 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6312895e  
IMMEDIATE SOURCE:  
CLONE: 989953

US-09-088-641-2

Query Match 1.4%; Score 34; DB 1; Length 1711;  
Best Local Similarity 63.4%; Pred. No. 3.4;  
Matches 52; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1258 GCTCTACTTTCCTGGAGCAGCGTTGTACACGTATATCTTCAAGACCCAGTGAT 1317  
DB 1270 GCTGTACAGTGCCTGGGCAAGCAGTTTACCCCTGATCTCATCAGGAAGGCTCTGTGAT 1329

QY 1318 GCTGCTGATGAAGGTAGACACA 1339  
DB 1330 GCTGCAGGTGATGTAGACACA 1351

RESULT 14

US-09-088-641-2  
Sequence 2, Application US/09088641  
Patent No. 6312895  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Hawkins, Phillip R.  
TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ Version 1.5  
CURRENT APPLICATION NUMBER: US/09/088,641  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/749,903  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0163 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 5759812e  
IMMEDIATE SOURCE:  
CLONE: 989953

US-08-749-903-2

Query Match 1.4%; Score 34; DB 1; Length 1711;  
Best Local Similarity 63.4%; Pred. No. 3.4;  
Matches 52; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1258 GCTCTACTTTCCTGGAGCAGCGTTGTACACGTATATCTTCAAGACCCAGTGAT 1317  
DB 1270 GCTGTACAGTGCCTGGGCAAGCAGTTTACCCCTGATCTCATCAGGAAGGCTCTGTGAT 1329

QY 1318 GCTGCTGATGAAGGTAGACACA 1339  
DB 1330 GCTGCAGGTGATGTAGACACA 1351

RESULT 15

US-08-227-536-1/c  
Sequence 1, Application US/08227536  
Patent No. 5658784  
GENERAL INFORMATION:  
APPLICANT: Eckner, Richard  
APPLICANT: Ewen, Mark  
APPLICANT: Livingston, David  
TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION  
TITLE OF INVENTION: FACTOR P300 AND USES OF P300  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Weingarten, Schurigin, Gagnebin & Hayes  
STREET: Ten Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/08/227,536  
FILING DATE: 14-APR-1994  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Ph.D., Kathleen A.  
REGISTRATION NUMBER: 34,380  
REFERENCE/DOCKET NUMBER: DFCI-308XX

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-2290
; TELEFAX: (617) 451-0313
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9046 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1200..8441
;
US-08-227-536-1

Query Match      1.4%; Score 34; DB 1; Length 9046;
Best Local Similarity 61.1%; Pred. No. 11;
Matches 55; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 599 TCITTTGGACATTCAAATTCACAAAGTCACACTTGAGATGGTTAAAGCTGCAACTGAGA 658
Db 4235 TGTTCGGTAGATTCCATTTTACAGTCTTCCACTTTAGACTCTGAATATCTCCGGCTG 4176

Qy 659 CGAATTGCGCGCTTCTGTTCTACCACTTC 688
Db 4175 CGTATCTGCTGGTCTGGTTGATCCACTTC 4146
```

Search completed: January 10, 2004, 06:15:36  
Job time : 138.136 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 22:59:07 ; Search time 3337.8 Seconds  
(without alignments)  
18087.435 Million cell updates/sec

Title: US-09-935-290-3  
Perfect score: 2484  
Sequence: 1 atgagtaactgcactgac.....ttctgagttttgtgtgctg 2484

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1843.6	74.2	3305	11 AK047295	AK047295 Mus muscu
2	848.2	34.1	918	13 B0437820	B0437820 AGENCOURT
3	760.2	30.6	1043	12 B0062851	B0062851 AGENCOURT
4	730.6	29.4	1055	12 BQ057957	BQ057957 AGENCOURT

5	703.8	28.3	965	10	BF794257
6	666.8	26.8	968	12	BQ058377
7	605	24.4	687	9	AL598680
8	561.4	22.7	637	14	CB268766
9	561.4	22.6	753	13	BU703927
10	547.2	22.0	580	10	AW976326
11	530	21.3	988	10	BG034520
12	511.4	20.6	776	14	CD000106
13	496.8	20.0	550	13	BQ694822
14	487.6	19.6	690	12	BI691948
15	467.8	18.8	825	13	BU263739
16	445	17.9	589	14	CA889049
17	430.2	17.3	799	10	EG704442
18	420	16.9	859	12	BI738244
19	418.2	16.8	746	14	BY741039
20	404.4	16.3	602	9	AL603562
21	398	16.0	446	12	BM414848
22	388.2	15.6	564	9	AA277375
23	381	15.3	1137	12	BM471397
24	369	14.9	553	9	AL596527
25	368.6	14.8	400	14	R73257
26	366.4	14.8	646	12	BI100769
27	365	14.7	573	14	CB269030
28	362.8	14.6	491	14	CA885724
29	359	14.5	535	14	CA896337
30	347.2	14.0	519	14	CA887606
31	347.2	14.0	527	14	CD554026
32	343.4	13.8	347	14	CB265571
33	340.2	13.7	571	9	AA172653
34	335.4	13.5	1037	10	BF037399
35	331.8	13.4	656	9	AA769055
36	329.8	13.3	411	12	BM146906
37	324	13.0	419	12	BM710854
38	321.8	13.0	823	14	CD362496
39	320.4	12.9	1310	14	CD502509
40	314.8	12.7	889	10	BF783411
41	314.6	12.7	751	14	CA384438
42	311.2	12.5	465	9	AA681924
43	306.6	12.3	713	13	BU444448
44	302.2	12.2	400	12	BG993243
45	300.8	12.1	635	12	BU063511

ALIGNMENTS

RESULT 1  
AK047295  
LOCUS  
DEFINITION  
Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:B930046K04 product:glycerol-3-phosphate acyltransferase, mitochondrial, full insert sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AK047295  
AK047295.1 GI:26092071  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
1  
Carninci,P. and Hayashizaki,Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
9279253  
10349636  
2

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
1  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

REFERENCE AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL MEDLINE PUBMED	Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861
REFERENCE AUTHORS	4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaïdo, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staab, J., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohtsuki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection
JOURNAL MEDLINE PUBMED	Nature 409 (6821), 685-690 (2001) 21085660 11217851
REFERENCE AUTHORS	5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
JOURNAL MEDLINE PUBMED	6 (bases 1 to 3305) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayaashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, K., Saiton, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
JOURNAL	COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ URL: http://Location/Qualifiers
FEATURES	

source	1..3305 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:B930046K04" /db_xref="taxon:10090" /clone="B930046K04" /tissue_type="cerebellum" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="10 days neonate" 156..2635 /note="glycerol-3-phosphate acyltransferase, mitochondrial (MGDI) [MGI:109162, GB NM_008149, evidence: BLASTN, 99%, match=2883] putative" 3281..3286 /note="putative" 3305 /note="putative" BASE COUNT 819 a 861 c 832 g 793 t ORIGIN	Query Match 74.2%; Score 1843.6; DB 11; Length 3305; Best Local Similarity 84.5%; Pred. No. 0; Matches 2098; Conservative 0; Mismatches 379; Indels 7; Gaps 2; QY 1 ATGGATGATTCGCACTGACCCCTTGGTCAATAGATGTTCTTATCTGCCACATTCATCA 60 DB 156 ATGGAGAGCTTTCAGTGAAGTGGCAATAGACGTTTCTTATCTGCCAGTTTCATCG 215 QY 61 GAATACAGTGTGGTTCGATGAAGCACACCAAGTGGGAGTGGGGTGGTGGTGGTGGT 120 DB 216 GAATACAGCTTGGCCGATGAACACACAGTGGGAGTGGGTTGAGTGGTGGTCAAA 275 QY 121 CCCACGCTTCAGATCTGCAATCTTAAATGGAAGAAAGCTTATGATGCGAAAGG 180 DB 276 CCTACCTTCTCAGATCTGCAACACTGAAATGGAAGAGAGCTTATGAGCGGAAGAG 335 QY 181 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCCAGAGCTGGGACAAAATTTTCAAC 240 DB 336 CCCTTCGTGGGAAGGTGCTGCTATTCTCGACCGCACAGAGCTGGGAAAGGTTTTTCAAC 395 QY 241 CCCAGTATCCCTCTTTTGGGTTTGGGGAATGTTATTTATATCAATGAAACTCACACAAG 300 DB 396 CCCAGTATCCCATCTCTGGGTTTGGGAATGTTATTTATCAATGAACGACACAGAAG 455 QY 301 CACCGCGGATGCTTGCAGAGACGCTTCTTACGTTTCTTATTTATCAAGAGCGAGATGTG 360 DB 456 CACAGAGGATGCTGCGCAGAGCGGTGCTTACATCTCTTTTTCAGAGCGAGACGTC 515 QY 361 CATAGGGCATGTTTGCACCACTGCTGCTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 420 DB 516 CATAGGGCATGTTTGCACCACTGCTGCTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 575 QY 421 GAGGCAATTCAGAGATGCTGCTGCTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480 DB 576 GAGGCAATTCAGAGATGCTGCTGCTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 635 QY 481 AAAGCGTTTAAACAAAGTGAAGAAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGAAG 540 DB 636 AAAGCGATCCAGAGGTTGAAGAAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGAAG 695 QY 541 GTCTACCGGCAATGATCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600 DB 696 GTCTACCGGCAATGATCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755 QY 601 TTTTGAACATTTCAAAATTCACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660 DB 756 TTTTGAACATTTCAAAATTCACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 815 QY 661 AATTTGCCGCTTCTGTTTCTACAGTTTATAGATCCCATATTTATGATCTGCTGCTGCTGCT 720 DB 816 AACCTGCCGCTTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 875
--------	---	--



http://image.llnl.gov									
Plate: L1AM13513 row: k column: 01									
High quality sequence stop: 657.									
FEATURES									
source									
1. .918									
/organism="Homo sapiens"									
/mol_type="mRNA"									
/db_xref="taxon:9606"									
/clone="IMAGE:6161400"									
/tissue_type="melanotic melanoma"									
/lab_host="DH10B (phage-resistant)"									
/clone_lib="NIH MGC_72"									
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 2 kb. Library constructed by Life Technologies."									
BASE COUNT 244 a 223 c 221 g 224 t 6 others									
ORIGIN									
Query Match 34.1%; Score 848.2; DB 13; Length 918;									
Best Local Similarity 98.7%; Pred. No. 1.1e-227;									
Matches 906; Conservative 0; Mismatches 6; Gaps 6;									
Qy	1291	GCTATACTTCTTCAAGACCCAGTGTGCTGCTGATGAAGTAGACACGTCATTAAAT	1350						
Db	1	GCTATACTTCTTCAAGACCCAGTGTGCTGCTGATGAAGTAGACACGTCATTAAAT	60						
Qy	1351	GAGTCCAGAAATGCAACAGATGAATCCTTCAGGAGGAGTTGATTGC AAAATCTGGCTGAG	1410						
Db	61	GAGTCCAGAAATGCAACAGATGAATCCTTCAGGAGGAGTTGATTGC AAAATCTGGCTGAG	120						
Qy	1411	CATATTCTATTCACTGTAGCAAGTCTGTGCCATTATGTCCACACACATTTGGCTTGC	1470						
Db	121	CATATTCTATTCACTGTAGCAAGTCTGTGCCATTATGTCCACACACATTTGGCTTGC	180						
Qy	1471	CTGCTCCTCTACAGACACAGCAGGGAATTGATCTCTCCACATTTGGTTCGAAGACTTCTTT	1530						
Db	181	CTGCTCCTCTACAGACACAGCAGGGAATTGATCTCTCCACATTTGGTTCGAAGACTTCTTT	240						
Qy	1531	GTGATGAAGAGGGAAGTCTTGGCTCGTGATTTTTGACCTGGGGTTCTCAGGAAATTCAGAA	1590						
Db	241	GTGATGAAGAGGGAAGTCTTGGCTCGTGATTTTTGACCTGGGGTTCTCAGGAAATTCAGAA	300						
Qy	1591	GATGTAGTAATGCAATGCCATACAGCTGCTGGGAAATTTGTGTCACAATCACCCACACTAGC	1650						
Db	301	GATGTAGTAATGCAATGCCATACAGCTGCTGGGAAATTTGTGTCACAATCACCCACACTAGC	360						
Qy	1651	AGGAACGATGAGTTTTTTATCACCCCAGACACAACTGTCCCATCAGTCTTCGAACTCAAC	1710						
Db	361	AGGAACGATGAGTTTTTTATCACCCCAGACACAACTGTCCCATCAGTCTTCGAACTCAAC	420						
Qy	1711	TTCTACAGCAATGGGGTACTTCTATGCTTTATCATGGAGGCCATCATAGCTTGCAGCCTT	1770						
Db	421	TTCTACAGCAATGGGGTACTTCTATGCTTTATCATGGAGGCCATCATAGCTTGCAGCCTT	480						
Qy	1771	TATGCAGTCTCTGAACAAGAGGGGACTGGGGGTCCCACTAGCACCCCACTTAACCTGATC	1830						
Db	481	TATGCAGTCTCTGAACAAGAGGGGACTGGGGGTCCCACTAGCACCCCACTTAACCTGATC	540						
Qy	1831	AGCCAGGACAGCTGGTTCGCGAAGGGGCCAGCCTGTGCTACTCTCTCCAATGAAGGC	1890						
Db	541	AGCCAGGACAGCTGGTTCGCGAAGGGGCCAGCCTGTGCTACTCTCTCCAATGAAGGC	600						
Qy	1891	ACCATCTCACTGCTTCGCCAGACATTTTACCAAGTCTGCCATGAACAGATGAGGAAGTTT	1950						
Db	601	ACCATCTCACTGCTTCGCCAGACATTTTACCAAGTCTGCCATGAACAGATGAGGAAGTTT	660						
Qy	1951	ATCCAGTATGGCATTTCTTACAGTGGCAGACCAATGACCAAGG-AAAGATATCACTCTCTAG	2009						
Db	661	ATCCAGTATGGCATTTCTTACAGTGGCAGACCAATGACCAAGGAAAGATATCACTCTCTAG	720						
Qy	2010	TCTTGTCTGAGCAGCAGTGGGACAAAGAAAGCTTTCAGAAAGCTTTT-GTCTTGGAGAAAGTATG	2068						





QY	601	TTTTGGAACTTCAAAATTCACAAAGGTCACATTGTAGATGGTTAAAGCTGCAACTGAGAAGC	660
Db	770	TTTTGGGACATTCAAA-TCACAAAGGTCACACTGAGATGGGTAAAGCTGCAACTGAGAAG	828
QY	661	AATTGCCGCTCTGTTTCTTACCAGTTTCATAGATCCCATATTGACTATCTGCT-GCTCAC	719
Db	829	AATTGCCGCTCTGTTTCTTACCAGTTTCATAGATCCCATATTGACTATCTGCTGCTCAC	888
QY	720	TTCATTTCTCTTGCCATAACATCAAAGCACCACCATACATTGCTTCAGGCATAATCTCAA	779
Db	889	TTTCCITCTCTCGCCATTAACTCAAGGCCCTTCATTGCTTCAGGCATAATCTCAA	948
QY	780	CATCCCCAAT-CTTCAGTACCTTGAT--CCATAGCTTTGGGGGCTTC--TTCAFACGACGA	834
Db	949	CATCCCCAATCTTCAGTACCTTGAT--CCATAGCTTTGGGGGCTTC--TTCAFACGACGA	1008
QY	835	AGCTCGATGAACACCCAGATGGACGAAAGATG	868
Db	1009	AGGCTCGGTAAACCCCAGTGGGACGGAAG	1042
RESULT 5	BF794257	60225455F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:433897 5', mRNA sequence.	EST 12-JAN-2001
LOCUS	BF794257	BF794257.1 GI:12099311	
DEFINITION	ACCESSION	BF794257	
VERSION	KEYWORDS	EST.	
SOURCE	ORGANISM	Homo sapiens (human)	
REFERENCE	AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	JOURNAL	NTH-MGC http://mgc.nci.nih.gov/.	
COMMENT		Unpublished Contact: Robert Strausberg, Ph.D. Email: rgs@nih.gov Tissue Procurement: Louis Staudt, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM9948 row: m column: 14 High quality sequence stop: 696. Location/Qualifiers 1. 965 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:433897" /tissue_type="lymphoma, cell line" /lab_host="DHIOB (phage-resistant)" /clone_lib="NIH_MGC_85" /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."	(MGC)
FEATURES	source	273 a 222 c 226 g 244 t	
BASE COUNT	ORIGIN	Query Match 28.3%; Score 703.8; DB 10; Length 965; Best Local Similarity 97.6%; Pred No. 6.6e-187; Matches 758; Conservative 0; Mismatches 12; Indels 7; Gaps 4;	
QY	964	GGA AAAACCTTGTGTCTCGG CAGGACTTTTGTGCTGTGGTAGATACCTGTCTTACC	1023
Db	1	GGA AAAACCTTGTGTCTCGG CAGGACTTTTGTGCTGTGGTAGATACCTGTCTTACC	60

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCW2067 row: i column: 02

High quality sequence stop: 722.

#### FEATURES

Location/Qualifiers  
1..968  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5814217"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."  
276 a 214 c 216 g 262 t

#### BASE COUNT

ORIGIN  
Query Match 26.8%; Score 666.8; DB 12; Length 968;  
Best Local Similarity 96.8%; Pred. No. 1.9e-176;  
Matches 724; Conservative 0; Mismatches 17; Indels 7; Gaps 4;  
Qy 1 ATGATGAATCTGCACCTGCTGGTACATAGATGTTCTTATCTGCCACATTCATCA 60  
Db |||||  
Qy 170 ATGATGAATCTGCACCTGCTGGTACATAGATGTTCTTATCTGCCACATTCATCA 229  
Db |||||  
Qy 61 GAATACAGTGTGTCGATGTAAGCACACAAAGTGAAGTGGGTGAGTGGCTTTAGA 120  
Db |||||  
Qy 230 GAATACAGTGTGTCGATGTAAGCACACAAAGTGAAGTGGGTGAGTGGCTTTAGA 289  
Db |||||  
Qy 121 CCCACGCTTTCAGATCTGCACTTTAAATGGAAGAAAGCCTAATGAGTCGGAAGG 180  
Db |||||  
Qy 290 CCCACGCTTTCAGATCTGCACTTTAAATGGAAGAAAGCCTAATGAGTCGGAAGG 349  
Db |||||  
Qy 181 CCATTTGTTGAAGATGTTGTTACTCCGCTGCTCCAGAGCTGGCAAAATTTTCAAC 240  
Db |||||  
Qy 350 CCATTTGTTGAAGATGTTGTTACTCCGCTGCTCCAGAGCTGGCAAAATTTTCAAC 409  
Db |||||  
Qy 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAACCTCACACA 300  
Db |||||  
Qy 410 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAACCTCACACA 469  
Db |||||  
Qy 301 CACCGGATGGCTTGCAGACGCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTG 360  
Db |||||  
Qy 470 CACCGGATGGCTTGCAGACGCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTG 529  
Db |||||  
Qy 361 CATAGGCGATGTTGGCCACCAATGCTGCTGAAATGCTGCAAGACAGTAGATGACAA 420  
Db |||||  
Qy 530 CATAGGCGATGTTGGCCACCAATGCTGCTGAAATGCTGCAAGACAGTAGATGACAA 589  
Db |||||  
Qy 421 GAGCAATTTGAGAGTGGCTGCTGAATTAACCTGATGTTCTGCCAGCAGCAATCA 480  
Db |||||  
Qy 590 GAGCAATTTGAGAGTGGCTGCTGAATTAACCTGATGTTCTGCTCAGCAGCAATCA 649  
Db |||||  
Qy 481 AAAGCGTTAAAGAGTGAAGAAAGAGCTTAAAGGATTTCTTCAAGAAATGTTGCCACT 540  
Db |||||  
Qy 650 AAAGCGTTAAAGAGTGAAGAAAGAGCTTAAAGGATTTCTTCAAGAAATGTTGCCACT 709  
Db |||||  
Qy 541 GTCTCAGCGGAATGATCAGACTGCTGGGTGGGTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db |||||  
Qy 710 GTCTCAGCGGAATGATCAGACTGCTGGGTGGGTGCTGCTGCTGCTGCTGCTGCTGCT 769  
Db |||||  
Qy 601 TTTTGAACATTCATTAATTCACAAAGTCACTTCAGATGGTTAAAGCTGCAACT--GAG 657  
Db |||||  
Qy 770 TTTTGAACATTCATTAATTCACAAAGTCACTTCAGATGGTTAAAGCTGCAACTGGAGAC 829  
Db |||||

Qy 658 ACGAATTGCGCTTCTGTTTCTACCAATT--CATAGATCCCATATTTGACTAT-CTGCTG 714  
Db |||||  
Qy 830 GAAATTTGCCGCTTGGTTTCTACCAATTTCATAGATCCCATATTTGACTATCTGCTG 889  
Db |||||  
Qy 715 CTCAC-TTTCAATCTCTTCGCCATAAC 741  
Db |||||  
Qy 890 CTCAC-TTTCAATCTCTTCGCCATAAC 917  
Db |||||  
RESULT 7  
LOCUS AL598680 687 bp mRNA linear EST 14-AUG-2001  
DEFINITION DKFZp313P1521\_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone  
ACCESSION DKFZp313P1521 5', mRNA sequence.  
VERSION AL598680  
KEYWORDS EST.  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Poustka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann, S.).  
EST (Poustka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann  
S.).  
Unpublished  
Contact: Poustka A.J.  
Department Lehrach  
Max-Planck-Institute for Molecular Genetics  
Innessstrasse 73, 14195 Berlin, Germany  
Tel: +49-30-84131623  
Fax: +49-30-84131128  
Email: poustka@mpg-berlin-dahlem.mpg.de  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No sl sequence available.  
This clone (DKFZp313P1521) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1..687  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp313P1521"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="313 (synonym: hlcc2)"  
/note="vector: pTriplex2; Site 1: SfiIA; Site 2: SfiIB;  
cDNA-collection"  
BASE COUNT 202 a 139 c 168 g 178 t  
ORIGIN  
Query Match 24.4%; Score 605; DB 9; Length 687;  
Best Local Similarity 100.0%; Pred. No. 4.5e-159;  
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1880 CCAATGAAGCACCATCTCACTGCTTGCAGACATTTTACCAGTTCGCCATGAACAG 1939  
Db |||||  
Qy 1940 TAGAAGATTTATCCAGTATGGCATTTTACAGTGGCAGAGCAGTACACAGAGATA 1999  
Db |||||  
Qy 61 TAGAAGATTTATCCAGTATGGCATTTTACAGTGGCAGAGCAGTACACAGAGATA 120  
Db |||||  
Qy 2000 TCAGTCTCTAGTCTGTCAGCAGCAGTGGGACAGAGAGCTTCAGAACCTTTGCTTGA 2059  
Db |||||  
Qy 121 TCAGTCTCTAGTCTGTCAGCAGCAGTGGGACAGAGAGCTTCAGAACCTTTGCTTGA 180  
Db |||||

```
QY 2060 GAAGTGATGAAGATGAAGACAGTGAAGTGGGGAGGAAACAGCGAGATGCTACCTGA 2119
Db 181 GAAGTGATGAAGATGAAGACAGTGAAGTGGGGAGGAAACAGCGAGATGCTACCTGA 240
QY 2120 AGGTGAGCAATCCAAAGGAGCAGCAGTGTATCACTTTTACAGAGTCTCTTGGGC 2179
Db 241 AGGTGAGCAATCCAAAGGAGCAGCAGTGTATCACTTTTACAGAGTCTCTTGGGC 300
QY 2180 CTTTGTGAGGCGCTACAGCTGCTGCGCATCTTTGTTTACAGTCTCAGTGTCTGCTG 2239
Db 301 CTTTGTGAGGCGCTACAGCTGCTGCGCATCTTTGTTTACAGTCTCAGTGTCTGCTG 360
QY 2240 CAGAACTCAGTATCTGCAAAAGTTCGACAAATACCTAATACGAAACAGAAAGAAATG 2299
Db 361 CAGAACTCAGTATCTGCAAAAGTTCGACAAATACCTAATACGAAACAGAAAGAAATG 420
QY 2300 TTGCAGTATATGCTGAGAGTGCCACATATTTGCTTGTGAAGATGCTGTGAAATGTTTA 2359
Db 421 TTGCAGTATATGCTGAGAGTGCCACATATTTGCTTGTGAAGATGCTGTGAAATGTTTA 480
QY 2360 AGGATATGGGTTTCAAGGAGACCAACAAAGAGAGTGTCTGTTTAGACTGAGCA 2419
Db 481 AGGATATGGGTTTCAAGGAGACCAACAAAGAGAGTGTCTGTTTAGACTGAGCA 540
QY 2420 GCACATTTTCTACCTCAATGCAACCCGACAAACTTCTAGATATATTTCTGAGTTTGTGG 2479
Db 541 GCACATTTTCTACCTCAATGCAACCCGACAAACTTCTAGATATATTTCTGAGTTTGTGG 600
QY 2480 TGCTG 2484
Db 601 TGCTG 605

RESULT 8
CB268766
LOCUS
DEFINITION
1007672 Human Fat Cell 5'-Stretch plus cdna Library Homo sapiens
EST.
CB268766
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 637)
Yang,R.-Z., Shuldiner,A. and Gong,D.-W.
EST analysis of human adipose gene expression
Unpublished
Contact: Gong Da-Wei
Division of Endocrinology, Diabetes and Nutrition
University of Maryland
660 Redwood St, HH497, Baltimore, MD 21201, USA
Tel: 410 706 1672
Fax: 410 706 1622
Email: dgong@medicine.umaryland.edu
PCR Primers
FORWARD: CTCGGGAAGCGGCCATTGTGTGGT
BACKWARD: AATACGACTCACTATAGGCGCAATTGG
Seq primer: GTTGGTACCGGGAATTC.
FEATURES
source
1..637
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="Male and Female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch plus cdna Library"
/note="Vector: lambda triplex"
173 a 145 c 131 g 184 t 4 others
Query Match 22.7%; Score 564.2; DB 14; Length 637;
```

```
Best Local Similarity 98.4%; Pred. No. 1.5e-147;
Matches 599; Conservative 0; Mismatches 7; Indels 3; Gaps 3;
QY 662 ATTTCGCGCTCTGTTTCTACAGTTCATAGATCCCATATGACTATCTGCTGCACATT 721
Db 1 ATTTCGCGCTCTGTTTCTACAGTTCATAGATCCCATATGACTATCTGCTGCACATT 60
QY 722 TCATTCTCTTCGCCATAACATCAAGCACCACATATGCTTCAGGCAATAATCTCAACA 781
Db 61 TCATTCTCTTCGCCATAACATCAAGCACCACATATGCTTCAGGCAATAATCTCAACA 120
QY 782 TCCCAATCTTCAGTACCTTGTATCCATAAGCTTGGGGGCTTTTTCATACGACGAGCTCG 841
Db 121 TCCCAATCTTCAGTACCTTGTATCCATAAGCTTGGGGGCTTTTTCATACGACGAGCTCG 180
QY 842 ATGAACACACAGATGGACGGAAGATGTTCTATAGAGCTTGTCCATGGGCATATAG 901
Db 181 ATGAACACACAGATGGACGGAAGATGTTCTATAGAGCTTGTCCATGGGCATATAG 240
QY 902 TTGAATTTACTTCGACAGCAATTTCTTGAGATCTTCTCGAAGGCACACGTTCTAGGA 961
Db 241 TTGAATTTACTTCGACAGCAATTTCTTGAGATCTTCTCGAAGGCACACGTTCTAGGA 300
QY 962 GTGGAAAAACCTCTTGTGCTCGGCAGGACTTTTGTGAGTGTGTGTGTGTGTGTGTGT 1021
Db 301 GTGGAAAAACCTCTTGTGCTCGGCAGGACTTTTGTGAGTGTGTGTGTGTGTGTGTGT 360
QY 1022 CCAATGTCTATCCAGACATCTTGATTAATACCTGTGTGGAATCTCCTATGATGCATATCG 1081
Db 361 CCAATGTCTATCCAGACATCTTGATTAATACCTGTGTGGAATCTCCTATGATGCATATCG 420
QY 1082 AAGGTCACCTACAAATGCTGAACAACTGGGCAACCTTAAGAAAGATGAGAGCTGTGGAGTG 1141
Db 421 AAGGTCACCTACAAATGCTGAACAACTGGGCAACCTTAAGAAAGATGAGAGCTGTGGAGTG 480
QY 1142 TAGCAAGAGGTGTTATT-AGAAATGTTACGAAAAAATACTGTTGTGTGTGTGTGTGTGT 1200
Db 481 TANCAAGAGGTGTTATTAAAGATGTTACGAAAAAATACTGTTGTGTGTGTGTGTGTGT 540
QY 1201 GCACAGCCATTTCCTTAAAGGAATATTAGAAAGCAAGTCAGAAACGGTGTGTGTGT 1260
Db 541 GCACAGCCATTTCCTTAAAGGAATATTAGAAAGCAAGTCAGAAACGGTGTGTGTGTGT 598
QY 1261 CTACTTTCC 1269
Db 599 CTACTTTCC 607

RESULT 9
BU703927
LOCUS
DEFINITION
UI-M-FOO-bzr-a-16-0-UI.r1 NIH_BMAP_FOO Mus musculus cdna clone
IMAGE: 6406167 5', mRNA sequence.
BU703927
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 753)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cdna Library preparation: Dr. M. Bento Soares, University of Iowa
cdna Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
```

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

#### FEATURES

Location/Qualifiers

1..753  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CS7BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE: 6406167"  
/tissue type="whole brain"  
/dev stage="embryo 12.5dpc"  
/lab host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH BMAP\_F00"  
/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TGAGAGACC. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 196 a 202 c 173 g 180 t 2 others

#### ORIGIN

Query Match 22.6%; Score 561.4; DB 13; Length 753;

Best Local Similarity 84.9%; Pred. No. 9.7e-147;

Matches 639; Conservative 0; Mismatches 113; Indels 1; Gaps 1;

QY 1033 CCAGACATCTTGAATAACCTGTTGGAAATCTCCATGATCGCATTCGAATTCGAAGGTCACCTAC 1092

Db 1 CCCGACATCTCTCGTACCCGCTGGGCATCTCGTATGATCGCATTAATCGAAGGTCACCTAC 60

QY 1093 AATGTGTAACAA-CTGGGCAACCTAAGCAAGATGAGCCCTGTGGAGTGTAGCAAGG 1151

Db 61 AATGGCAACAGTTGGGAAGCCCAAGAAACGAGAGCCCTGTGGAGTGTGGCGAGAGG 120

QY 1152 TGTATTAGAAATTTACCAAAAAAATATGTTGTTCGAGTGGATTTTGACAGCCATT 1211

Db 121 CGTTATCAGATGTCGCGAAATACTACGGTCTACGTCGAGTGGATTTTGACAGCCATT 180

QY 1212 TTCCTTAAGGAATATTTAGAAAGCAAAAGTCAGAAACCGGTGTCTGCTTACTTTCCCT 1271

Db 181 TTCCTTGAAGGAATATTTAGAAAGCCAGAGTCAGAAACCTGTATCTGCCCTTTCTCT 240

QY 1272 GGAGCAAGCTTGTACCAGCTATCTTCTTCAAGCCAGTATGATGTCGTGATGAAGG 1331

Db 241 GGAGCAAGCACTGTTTACCAAGGATCTTCTTCAAGCCAGTATGATGTCGTGATGAACA 300

QY 1332 TAGAGACACGTCCATTAATAGTCAGAAATGCAACAGATGAATCCCTTACCAAGAGGTT 1391

Db 301 TCAAGACCTATCCAGTACAGATCCAGAAACCCAGCAGACAGAGCTTCCAGCAAGGCT 360

QY 1392 GATTGCAAAATCTGGCTGAGCATAATTATTCCTGCTAGCAAGTCTGTGCCATTAATGTC 1451

Db 361 GATTGCAAACTGGCTGAGCATAATCTTCTTCAAGCCAGCAAGTCTGTGCCATTAATGTC 420

QY 1452 CACACATTTGGCTTGCCTCTCTTACAGACAGCAGGAGGAATGATCTCTCCAC 1511

Db 421 CACCACATTTGGCTTGCCTCTCTTACAGACAGCAGGAGGAATGATCTCTCTCCAC 480

QY 1512 ATTGGTCAAGACTTCTTTGTGTAAGAGAGAGTCTCTGCTCGTGAATTTGACCTGGG 1571

Db 481 CTTTGTGGAAGACTTCTTTGTGTAAGAGAGAGTCTCTGCTCGTGAATTTGACCTGGG 540

QY 1572 GTTCTCAGGAATTCAGAAAGATGTAGTAATGCATGCCATACAGCTGTCTGGGAAATTTGTG 1631

Db 541 CTTCTCCGGGAATTCAGAAGATGTCGTCATGCATGCTATTTCAGTCTTCTGGGAACTGTGT 600

QY 1632 CACAATCACCACACTAGCAGGAACGATGAGTTTTTATACACCCCCAGCAACACTGTCCC 1691

Db 601 CACAATCACCACACTAGCAGGAACGATGAGTTTTTATACACCCCCAGCAACACTGTCCC 660

QY 1692 ATCAGTCTTCGAACCTCAACTTCTACAGCAATGGGGTACTTTCATGCTTTATCATGGAGGC 1751

Db 661 GTCACTCTTGAACCTCAACTTCTACAGCAATGGGGTACTTTCATGCTTTATCATGGAGGC 720

QY 1752 CATCATAGCTTGCAGCCTTTATGCAAGTCTCGAA 1784

Db 721 CATCATAGCTTGCAGCCTTTATGCAAGTCTCGAA 753

#### RESULT 10

AW976326

LOCUS

DEFINITION

AW976326

ACCESSION

VERSION

AW976326.1

GI:8167552

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: John Quackenbush

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 3528

Fax: 301 838 0208

Email: johnq@tigr.org

Plate: 359

Seq primer: Reverse.

Location/Qualifiers

source

1..580

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone\_lib="MAGE resequences, MAGN"

/note="Vector: pBluescriptSKm"

174 a 119 c 138 g 149 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 555; Conservative

0; Mismatches 13; Indels

0; Gaps

0;

QY 1889 GCACATCTCAGTCCTTGGCCAGACATTTTACCAAGTCTGCCATGGAACAGTGAAGT 1948

Db 1 GCTCATCTCAGTCCTTGGCCAGACATTTTACCAAGTCTGCCATGGAACAGTGAAGT 60

QY 1949 TTATCCAGTATGGCAATTTTACAGTGGCAGACGACGATGACCCAGGAAGATATCAGTCCTA 2008

Db 61 TTATCCAGTATGGCAATTTTACAGTGGCAGACGACGATGACCCAGGAAGATATCAGTCCTA 120

QY 2009 GTCTTGTGAGCAGCAGTGGGACAAAGAGCTTCCAGAACCTTTGCTTTGGAGAGTGAATG 2068

Db 121 GTCTTGTGAGCAGCAGTGGGACAAAGAGCTTCCAGAACCTTTGCTTTGGAGAGTGAATG 180

QY 2069 AAGAAGATGAAGCAGTGAATTTGGGGAGGACAGCAGATGCTACCTGAGAGTGAAGC 2128

Db 181 AAGAAGATGAAGCAGTGAATTTGGGGAGGACAGCAGATGCTACCTGAGAGTGAAGC 240

QY 2129 AATCCAGGAGCAGCAGCAGTGTATACCTTCTTACAGAGACTCTCTTGGGCTTTGCTGG 2188

```
Db 241 AATCCAGGAGCACCAGCAGTTTATCACCTTCTTACAGAGACTCCTTTGGSCCTTCTGCTGG 300
Qy 2189 AGCCCTACAGCTCTGCTGCATCTTTGTTTCACAACTTCACTTCACTGCTGCTGTTCCAGAACCTG 2248
Db 301 AGCCCTACAGCTCTGCTGCATCTTTGTTTCACAACTTCACTTCACTGCTGCTGTTCCAGAACCTG 360
Qy 2249 AGTATCTGCAAAAGTTGCACAAATACCTAATAACCAAGACAGAAAGAAATGTTGCAAGTAT 2308
Db 361 AGTATCTGCAAAAGTTGCACAAATACCTAATAACCAAGACAGAAAGAAATGTTGCAAGTAT 420
Qy 2309 ATGCTGAGAGTGCACATATTTGCTTGTGAAGAAATGCTGGAATAATGTTTAAAGATATTG 2368
Db 421 ATGCTGAGAGTGCACATATTTGCTTGTGAAGAAATGCTGGAATAATGTTTAAAGATATTG 480
Qy 2369 GGGTTTTCAAGGAGACCAACAAAGAGAGTGTCTGTTTAGAACTGAGCAGCAGCACTTTTC 2428
Db 481 GGGTTTTCAAGGAGACCAACAAAGAGAGGGGCTGTTTATAGAACTGGGAGCAGCACTTTT 540
Qy 2429 TACTCAATGCAACCGACAAACTTCT 2456
Db 541 TACTCAATGCAACCGACAAACTTCT 568

RESULT 11
LOCUS BG034520 988 bp mRNA linear EST 24-JAN-2001
DEFINITION 602302638F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4404225 5',
mRNA sequence.
ACCESSION BG034520
VERSION BG034520.1 GI:12427920
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. 988
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgs.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM10115 row: k column: 10
High quality sequence stop: 604.
Location/Qualifiers
1. 988
/organism="Homo sapiens"
/mol type="mRNA"
/db xref="taxon:9606"
/clone="IMAGE:4404225"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 87"
/notes="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."
BASE COUNT 269 a 250 c 248 g 221 t
ORIGIN

Query Match 21.3%; Score 530; DB 10; Length 988;
Best Local Similarity 93.8%; Pred. No. 8.1e-138;
Matches 619; Conservative 0; Mismatches 30; Indels 11; Gaps 6;
Qy 1630 GTCAATATCAACCCACTAGCAGGACGATGATGTTTTTATCACCCCGACACAACTGTC 1689
```

```
Db 1 GTCAAAATCACCCACACTAGCAGGAATGATGATGTTTTTATCACCCCGACACAACTGTC 60
Qy 1690 CCATCAGTCTTCGAACCTCAACTTCTACAGCAATGGGTACTTTCATGTCCTTTATCATGGAG 1749
Db 61 CCATCAGTCTTCGAACCTCAACTTCTACAGCAATGGGTACTTTCATGTCCTTTATCATGGAG 120
Qy 1750 GCATCATAGCTTTCAGAGCTTTATCAGTTCCTGAACAGAGGGGACTGGGGGCTCCACT 1809
Db 121 GCATCATAGCTTTCAGAGCTTTATGAGTTCCTGAACAGAGGGGACTGGGGGCTCCACT 180
Qy 1810 AGCACCCCACTTAACCTGATCAGCAGGAGAGCTGGTGGGAAGCGGCCAGCCTGTGC 1869
Db 181 AGCACCCCACTTAACCTGATCAGCAGGAGAGCTGGTGGGAAGCGGCCAGCCTGTGC 240
Qy 1870 TACCTTCTCTCAATGAAGGACCAAT-CTCAGTCTGCTTCCAGACATTTTACCAAGTCTG 1928
Db 241 TACCTTCTCTCAATGAAGGACCAATCCTCACTGCTTCCAGACATTTTACCAAGTCTG 300
Qy 1929 CCATGAACAGTAGGAAGTTTTATCCAGTATGGCAATTTTACAGTGGCAGACGATGA 1988
Db 301 CCATGAACAGTAGGAAGTTTTATCCAGTATGGCAATTTTACAGTGGCAGACGATGA 360
Qy 1989 CCAGGAAGATATCAGTCTCTAGTCTTCTGAGCAGCAGTGGGACAAAGCTTCCAGAAC 2048
Db 361 CCAGGAAGATATCAGTCTCTAGTCTTCTGAGCAGCAGTGGGACAAAGCTTCCAGAAC 420
Qy 2049 TTTGCTTTCGAGAGTGTATGAAGATGAACAGACAGTGTGTTGGGAGGACACGCGAGA 2108
Db 421 TTTGCTTTCGAGAGTGTATGAAGATGAACAGACAGTGTGTTGGGAGGACACGCGAGA 480
Qy 2109 TTGCTACCTGAAGGTGAGCAAT-CCAAGGACACAG-CAGTTCATCACTTCTTACAG 2166
Db 481 TTGCTACCTGAAGGTGAGCAATCCAAGGACACAGCAGCAGTTCATCACTTCTTACAG 540
Qy 2167 AGA----CTCCTTGGGCTTTCCTGAGGCTTACA--GCTCTGCTGCCATCTTTGTTAC 2220
Db 541 AGAGGACTCTTTTGGGCTTTCCTGAGGCTTACAAGCTTTCCTGCTCATCTTTGTTAC 600
Qy 2221 AACTT--CAGTGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAA 2278
Db 601 AACTTTCAGTGTCTCTGTTCCAGAACCTTGAATCTGCAAAAGTTGCAAAATACCTTAA 660

RESULT 12
LOCUS CD000106 776 bp mRNA linear EST 01-MAY-2003
DEFINITION AGENCOURT_13642874 NIH_MGC_186 Homo sapiens cDNA clone
IMAGE:30322827 5', mRNA sequence.
ACCESSION CD000106
VERSION CD000106.1 GI:30294625
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. 776
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgs.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
High quality sequence stop: 549.
Location/Qualifiers
1. 776
FEATURES
source
```



```

RESULT 14
BI691948          690 bp  mRNA  linear  EST 18-SEP-2001
LOCUS             60311882F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5351781 5',
DEFINITION        mRNA sequence.
ACCESSION         BI691948
VERSION           BI691948.1 GI:15654577
KEYWORDS          EST.
SOURCE            Mus musculus (house mouse)
ORGANISM          Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE         NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS           National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE             Unpublished
JOURNAL           Contact: Robert Strausberg, Ph.D.
COMMENT           Email: cgapbs-i@mail.nih.gov
                  Tissue Procurement: Jeffrey Green M.D.
                  CDNA Library Preparation: Life Technologies, Inc.
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM1894 row: d column: 22
                  High quality sequence stop: 687.
FEATURES
    source
        1. 690
            Location/Qualifiers
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:5351781"
                /sex="female, virgin"
                /tissue_type="infiltrating ductal carcinoma"
                /dev_stage="5 months"
                /lab_host="DH10B"
                /clone_lib="NCI CGAP Mam6"
                /notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Salt;
                Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                Library constructed by Life Technologies. Investigator
                providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      173 a 172 c 180 g 165 t
ORIGIN
Query Match      19.6%; Score 487.6; DB 12; Length 690;
Best Local Similarity 85.5%; Pred. No. 6.5e-126;
Matches 590; Conservative 0; Mismatches 94; Indels 6; Gaps 4;
QY 1493 AGGAATTGATCTTCACATTTGGTTCGAAGACTCTTTGTGATGAAGAGAGAGTCTCTGG 1552
Db 1 AGGGAATCCATCTCTCCACGCT-GTGGAGACTTCTTTGTGATGAAGAGAGAGTCTCTAG 59
QY 1553 CTGCTGATTTTGACCTGGGGTCTCAGGAATTCAGAGATGTAGTAATCATGCCATAC 1612
Db 60 CTCGCGATTTTCGACCTAGGCTTCTCCGGGAATTCAGAGATGTCTGTCATGCTATTC 119
QY 1613 AGCTGCTGGGAATTTGTGTACAAATACCCACACTAGCAGGAACGATGATTTTTTATCA 1672
Db 120 AGTTCTGGGGAATGTGTACAAATACCCACACTAGCAGGAACGATGATTTTTTATTA 179
QY 1673 CCCCCAGCACAATGTGCCCATCAGTCTTCGAATCTCAAC-TTCTACAGCAATGGGTACTT 1731
Db 180 CTCGCCAGCACAATGTGCCCATCAGTCTTTGAATCAACCTTCTACAGCAATGGGTACTT 239
QY 1732 CATGCTTTTATCATGGAGGCATCATAGCTTGAGCTTTATGCAGTCTTGACAGAGG 1791
Db 240 CATGTGTTTCATGGAAGCCATCATAGCTTGAGCATCTATGCAGTCTTGAATAGAGG 299
QY 1792 GGACTGGGGGGTCCCACTAGCACCCACCTAACTGATCAGCCAGGAGCAGCTGGTGGCG 1851

```

```

Db 300 TGCTCTGGAGGGTCCGCTGGAGGCTCGGCAACCTGATCAGCAGGAGCAGCTGGTGAGG 359
QY 1852 AAGGGCGGCAGCTGTGCTACCTTCTCTCAATGAGGACCATCTCTCACTGCTTGGCCAG 1911
Db 360 AAGGGCGGCAGCTGTGCTACCTTCTCTCAATGAGGACCATCTCTCACTGCTTGGCCAG 419
QY 1912 ACATTTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTCTTACA 1971
Db 420 ACTTTTACCAAGTCTGTCTGAGACAGTTGGCAAGTTTCATCCAGTATGGCATTTCTTACA 479
QY 1972 GTGGCAGACGATGACCCAGGAAGATATCATGCTTCTAGTCTTGTGCTGAGCAGAGTGGAC 2031
Db 480 GTGGCAGACGATGACCCAGGAAGATGTCAGTCTCTGGCTTGCAGAACAGCAGAGTGGAC 539
QY 2032 AGAAGCTTCCAGAACCTTTGTCTTGGAGAGTGTATGAGAGATGAGAGACAGTGTACTTT 2091
Db 540 AGAAGCTTCCGGAAC---TGAACCTGGAGAGTGCAGGAGAGATGAAGACAGTGTACTTT 596
QY 2092 GGGGAGGA-ACAGCGAGATTGCTACTGAAGGTGAGCCAAATCCAGGAGCAGCAGCAGTT 2150
Db 597 GGTGAGGCCAGCGAGATTGCTATCTCAAGGTGAGCCAGTCCAGGAGCAGCAGCAATT 656
QY 2151 TATCACTTCTTACAGAGACTCTTTGGGCC 2180
Db 657 CATCACTTCTTACAGAGCTTCTAGGTCC 686

RESULT 15
BU263739          825 bp  mRNA  linear  EST 26-NOV-2002
LOCUS             60381507F1 CSEQCHN52 Gallus gallus cDNA clone CHEST808n1 5', mRNA
DEFINITION        sequence.
ACCESSION         BU263739
VERSION           BU263739.1 GI:25533451
KEYWORDS          EST.
SOURCE            Gallus gallus (chicken)
ORGANISM          Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE         1 (bases 1 to 825)
AUTHORS           Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
                  Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
                  A Comprehensive Collection of Chicken cDNAs
                  Curr. Biol. 12 (22), 1965-1969 (2002)
TITLE             Gallus gallus
JOURNAL           MEDLINE
PUBMED            22335534
COMMENT           Contact: Simon Hubbard
                  Department of Biomolecular Sciences
                  University of Manchester Institute of Science and Technology (UMIST)
                  PO Box 88, Manchester, M60 1QD, UK
                  Tel: 01612008930
                  Fax: 01612360409
                  Email: Simon.Hubbard@umist.ac.uk.
FEATURES
    source
        1. 825
            Location/Qualifiers
                /organism="Gallus gallus"
                /mol_type="mRNA"
                /strain="White Leghorn, Hisex"
                /db_xref="taxon:9031"
                /clone="CHEST808n1"
                /dev_stage="22"
                /lab_host="DH10B"
                /clone_lib="CSEQCHN52"
                /note="Organ: limbs; Vector: pBluescript II KS(+); Site 1:
                EcoRI; Site 2: NotI; This normalized library was
                constructed from 1 million independent clones. cDNA
                synthesis was initiated using an oligo(dT) primer, using
                methylated C in the first strand synthesis reaction.
                Following this first strand reaction, double-stranded cDNA
                was blunted, ligated to NotI adapters, digested with EcoRI

```

, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 231 a 178 c 182 g 234 t  
ORIGIN

Query Match 18.8%; Score 467.8; DB 13; Length 825;  
Best Local Similarity 74.5%; Pred. No. 2.8e-120;  
Matches 615; Conservative 0; Mismatches 207; Indels 3; Gaps 2;

Qy 1001 TTGTGGTAGATACCTCTGTCTACCAATGTCATCCAGACATCTTGATATACCTGTGGAA 1060  
Db 1 TGGTAGTGGATGCTCTCTGTCAATGCTACTCTCTGATGCTCTAATATACCTGTGGAA 60

Qy 1061 TCTCTATGATCGCATATTCGAAGTCACTACATGGTGAACAACTGGGCAAACTTAAGA 1120  
Db 61 TCTCTACGATCGCATATTCGAAGTCACTATAACAGTGAACAGCTGGGCAAGCTTAAGA 120

Qy 1121 AGAATGAGAGCCTGTGGAGTGTAGCAAGAGTGTATTAGAACTTTACGAAAACACTATG 1180  
Db 121 AGAATGAAGTCTTTGGAGTATAGCAAGAGAGTCTCAGAACTGCTGGGAAGAAATTATG 180

Qy 1181 GTTGTGTCGGAGTGGATTTCGACAGCCATTTTCCTTAAAGAAATATTAGAAAGCCAAA 1240  
Db 181 GGTGTGTCAGATAGATTTCGACAAACCATTTTCCTTAAAGAAATATTAGAACCCAAA 240

Qy 1241 GTCAGAAACCGGTGTCCTCTACTTTCCCTGGAGCAAGCGTTGTTACCAGCTATACTTC 1300  
Db 241 GTCAGAAACCGTCTCTCTCTCTCTTCTTCTTGAACAAGCTTTGTTACCAGCTATACTTC 300

Qy 1301 CTTCAAGACCCAGTATGCTGCTGATGAAGTAGAGACACGTCCTAATTAATGAGTCCAGAA 1360  
Db 301 CATCGAGACCTAATGATATCTGGATGAAGTACTGAGGCTTCACTGCTTAACCTCAAGG 360

Qy 1361 ATGCAACAGATGAATCCCTACGAGGAGGTTGATTGCAAACTGGCTGAGCATATTTAT 1420  
Db 361 ATATAACAGTGAACCTTACAGAGAGAGCTGATAGCCAAATTTGGCTGAGCATATTTAT 420

Qy 1421 TCATGTGTAGCAAGTCTGTGCCATATATGTCACACACATTTGGCTTGCCTCTCTCT 1480  
Db 421 TCACTGTCTAACAAAGTCTGTGTGTGTCTACCCACATTTGTCCTGTTTGTCTGTGT 480

Qy 1481 ACAGACACAGGAGGAATTCATCTCCCATTTGGTCGAAGACTTCTTTGTGATGAAG 1540  
Db 481 ACAGACATAGACAGGGGACTGATCTTTCCAGGTTGGTGAAGATTTCTTCTCCATGAAG 540

Qy 1541 AGGAAGTCTCTGCTGATTTTGACCTGGGGTCTCAGGAAATTCAGAACTGTAGTAA 1600  
Db 541 AAGAGTCTTAGCCGCTGACTTCGACTTGGGAATTTTCAAGGAACTCAGATGATTTGCA 600

Qy 1601 TGCATGCCATACAGTCTGGGAAATTTGTCACAATCACCCACACTAGCAGGAACGATG 1660  
Db 601 TGCATGCCATCCACTTTGTTGGGAACTGTGTAAATATCAAAACACTAGTGAACAAAG 660

Qy 1661 AGTTTATTTAT -CACCCACAGACAACTGTCCTCATCAGTCTTCGAACTCAACTTCTACAGC 1719  
Db 661 AGTTTCTTATAAACTCCACAGACAGACTGCTGCTGCTTTTGAACCTCAATTTCTACAGC 720

Qy 1720 AATGGGTACTTCATGCTTTTATCATGAGGCCATCATAGCTTGA -GCCTTTATGCAG 1777  
Db 721 AATGGAATCTTCATGATTCATTAAAGAGGAGTCAATGGGCTTCAAGTCTTCATGCCAG 780

Qy 1778 TTCTGAACAAGAGGGGAGTGGGGGCTCCCACTAGCACCCCACTTA 1822  
Db 781 TTCAGATGAAGGTCCCGAAATGGTACCAATGGTGTCTTCTCCCA 825



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 18:05:10 ; Search time 425.997 Seconds  
(without alignments)  
15740.506 Million cell updates/sec

Title: US-09-935-290-3  
Perfect score: 2484  
Sequence: 1 atgagatgaatgcactgac.....ttctgagttttgtgtgtcgtg 2484

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq 19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2484	100.0	3003	24	ABK11094
2	2484	100.0	4031	23	ABV25313
3	2482.4	99.9	2487	24	ABK94821
4	2480.8	99.9	2487	25	AAL55475
5	2480.8	99.9	3210	24	ABV35221
6	2476	99.7	2684	25	ABV72233
7	2185.8	88.0	3273	24	ABZ12025
8	1879.2	75.7	2646	24	ABK63741

9	437	17.6	469	22	ABA08415	Human sn-glycerol-
10	342.2	13.8	429	25	ABX42219	Bovine EST associa
11	337.6	13.6	407	25	ABX43656	Bovine EST associa
c 12	302	12.2	418	25	ABX35577	Bovine EST associa
c 13	233	9.4	303	25	ABX47921	Bovine EST associa
14	209.6	8.4	264	25	ABX41074	Bovine EST associa
15	172.6	6.9	387	25	ABX42469	Bovine ORFX polynuc
16	161.8	6.5	231	24	ABN22191	Human TRNFR-7 CDNA
17	109	4.4	2682	24	ABV52820	Human TRNFR-7 CDNA
18	109	4.4	2755	24	ABV52822	Drosophila melanog
19	103.8	4.2	3439	23	ABL09555	Drosophila melanog
20	103.8	4.2	3439	23	ABL09555	Drosophila melanog
c 21	99.4	4.0	501	25	ABX38896	Bovine EST associa
22	53.8	2.2	2640	23	AAV87279	DNA encoding novel
23	49.2	2.0	65	24	ABN28887	Rat spliced transcr
24	47	1.9	2502	22	AAV02332	E. coli p15B DNA e
25	47	1.9	2502	22	AAV02332	Escherichia coli p
26	46.2	1.9	2154	23	AAV81936	DNA encoding novel
27	46.2	1.9	2154	23	AAV82407	DNA encoding novel
28	46.2	1.9	2154	23	AAV82517	DNA encoding novel
29	45.6	1.8	2175	23	ABL07919	Drosophila melanog
30	45	1.8	100848	22	AAV28552	Genomic fragment #
c 31	44.8	1.8	1830121	17	AAV24063	Haemophilus influe
32	43.6	1.8	1127	23	AAV81997	DNA encoding novel
33	43.6	1.8	1127	23	AAV82408	DNA encoding novel
c 34	43.6	1.8	4120	23	AAV90070	DNA encoding novel
35	41.2	1.7	185371	24	ABT10718	Human breast cance
36	39.8	1.6	21045	22	AAV26721	Human genomic DNA
37	39.8	1.6	21045	25	ABV74070	Human novel polynu
c 38	39	1.6	1673	23	AAV72799	DNA encoding novel
c 39	39	1.6	2114	21	AAV74225	Human secreted pro
c 40	39	1.6	3714	23	AAV73941	DNA encoding novel
41	38.6	1.6	4379	22	AAV82496	Human immune/haema
42	38.6	1.6	4794	22	AAV82497	Human immune/haema
43	38.6	1.6	4799	22	AAV82500	Human immune/haema
44	38.6	1.6	4807	22	AAV82498	Human immune/haema
c 45	38.6	1.6	1664976	19	AAV21209	Methanococcus jann

ALIGNMENTS

RESULT 1  
ABK11094  
ID ABK11094 standard; CDNA; 3003 BP.  
XX ABK11094;  
AC ABK11094;  
XX 18-JUN-2002 (first entry)  
DE CDNA encoding human acyltransferase, ACTR-1.  
XX Human; acyltransferase; ACTR-1; antidiabetic; anorectic;  
KW antiarteriosclerotic; cerebroprotective; anorectic; ophthalmological;  
KW cardiant; metabolic disorder; energy homeostasis disorder; diabetes;  
KW hyperglycaemia; hypercholesterolaemia; hyperlipoproteinaemia; stroke;  
KW hypertriglyceridaemia; hyperlipidaemia; atherosclerosis; obesity;  
KW reinnopathy; nephropathy; peripheral neuropathy; weight disorder;  
KW appetite regulation disorder; cachexia; anorexia; bulimia;  
KW cardiovascular disorder; gene; ss.  
XX Homo sapiens.

Key	Location/Qualifiers
CDS	341..2827
FT	/*tag= a
FT	/product= "Human acyltransferase ACTR-1"
FT	/note= "The coding region is specifically claimed in claim 1"
XX	W0200216592-A2.
XX	28-FEB-2002.



Db 1601 CTACTTTCCCTGGAGCAAGCGTTGTTACAGCTATACCTTCTTCAAGACCAGATGCT 1660  
Qy 1321 GCTGATGAAGGTAGAGACACGCTCCATTAATAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db 1661 GCTGATGAAGGTAGAGACACGCTCCATTAATAGTCCAGAAATGCAACAGATGAATCCCTA 1720  
Qy 1381 CGAAGGAGTTGATTCGAAATCTGCTGAGCATATTTCTATTACCTGCTAGCAAGTCTCTGT 1440  
Db 1721 CGAAGGAGTTGATTCGAAATCTGCTGAGCATATTTCTATTACCTGCTAGCAAGTCTCTGT 1780  
Qy 1441 GCATTATGTCACACACATTTGCTTGGCTTGGCTCTCTCTACAGACACAGCAGGGAAT 1500  
Db 1781 GCCATTATGTCACACACATTTGCTTGGCTTGGCTCTCTCTACAGACACAGCAGGGAAT 1840  
Qy 1501 GATCTCTCCACATTTGCTCGAAGACTTTCTTTGTGATGAAGAGAAAGTCTGCTGCTGAT 1560  
Db 1841 GATCTCTCCACATTTGCTCGAAGACTTTCTTTGTGATGAAGAGAAAGTCTGCTGCTGAT 1900  
Qy 1561 TTTGACCTGGGTTCTTCAGGAATTCAGAAAGTGTAGTAATGCAATGCCATACAGCTGCTG 1620  
Db 1901 TTTGACCTGGGTTCTTCAGGAATTCAGAAAGTGTAGTAATGCAATGCCATACAGCTGCTG 1960  
Qy 1621 GGAATTTGTCACAACTACCCACACTAGCAGGAACGATGATTTTTTTATCACCCCCAGC 1680  
Db 1961 GGAATTTGTCACAACTACCCACACTAGCAGGAACGATGATTTTTTTATCACCCCCAGC 2020  
Qy 1681 ACAACTGTCCCATCAGTCTTCCAACTCTCAAGCAATGGGGTACTTCAATGCTCTTT 1740  
Db 2021 ACAACTGTCCCATCAGTCTTCCAACTCTCAAGCAATGGGGTACTTCAATGCTCTTT 2080  
Qy 1741 ATCATGAGGGCCATCATAGTCTGAGCTTTATGCAATTTCTGAAGAGGGAAGTCTGGG 1800  
Db 2081 ATCATGAGGGCCATCATAGTCTGAGCTTTATGCAATTTCTGAAGAGGGAAGTCTGGG 2140  
Qy 1801 GGTCCACTAGCACCCCACTAACCTGATCAGCAGGAGCAGCTGGTGGAGAGCGGCC 1860  
Db 2141 GGTCCACTAGCACCCCACTAACCTGATCAGCAGGAGCAGCTGGTGGAGAGCGGCC 2200  
Qy 1861 AGCTGTGCTACTCTCTCCAAATGAAGGCACCATCTCACTGCCCTTGCACACATTTTAC 1920  
Db 2201 AGCTGTGCTACTCTCTCCAAATGAAGGCACCATCTCACTGCCCTTGCACACATTTTAC 2260  
Qy 1921 CAAGTCTGCCATGAACAGTAGGAAGTTTATCCAGTATGGCAATTTTACAGTGCCAGAG 1980  
Db 2261 CAAGTCTGCCATGAACAGTAGGAAGTTTATCCAGTATGGCAATTTTACAGTGCCAGAG 2320  
Qy 1981 CACGATGACCGAGAGATATCAGTCTCTGCTGAGCAGCAGTGGGACAGAGCTT 2040  
Db 2321 CACGATGACCGAGAGATATCAGTCTCTGCTGAGCAGCAGTGGGACAGAGCTT 2380  
Qy 2041 CCAGAACCTTTGCTTTGGAGAGTATGAAGAGATGAAGCAGTGAATTTGGGGAGGAA 2100  
Db 2381 CCAGAACCTTTGCTTTGGAGAGTATGAAGAGATGAAGCAGTGAATTTGGGGAGGAA 2440  
Qy 2101 CAGCAGATTTGCTACCTGAAGGTAGGCAATCCAAAGGAGCAGCAGCTTTTACCTTC 2160  
Db 2441 CAGCAGATTTGCTACCTGAAGGTAGGCAATCCAAAGGAGCAGCAGCTTTTACCTTC 2500  
Qy 2161 TTACAGACCTCTTGGGCTTTGCTGAGGCTTACAGCTCTGCTGCAATTTGTTTAC 2220  
Db 2501 TTACAGACCTCTTGGGCTTTGCTGAGGCTTACAGCTCTGCTGCAATTTGTTTAC 2560  
Qy 2221 AACTTCAGTGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTCGCAAAATACCTAATA 2280  
Db 2561 AACTTCAGTGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTCGCAAAATACCTAATA 2620  
Qy 2281 ACCAGACAGAAAGAAATTTGTCAGTATATGCTGAGAGTGGCCACATATTTGTTGAG 2340  
Db 2621 ACCAGACAGAAAGAAATTTGTCAGTATATGCTGAGAGTGGCCACATATTTGTTGAG 2680  
Qy 2341 AATGCTGTGAAATGTTTAAAGATATTTGGGTTTTCAAGGAGACCAAAACAAAGAGGTG 2400  
Db 2681 AATGCTGTGAAATGTTTAAAGATATTTGGGTTTTCAAGGAGACCAAAACAAAGAGGTG 2740

Qy 2401 TCTGTTTAGAAGTGAAGCAGCAGCTTTTCTACCTCAATGCAACCGCAAAATCTTAGAA 2460  
Db 2741 TCTGTTTAGAAGTGAAGCAGCAGCTTTTCTACCTCAATGCAACCGCAAAATCTTAGAA 2800  
Qy 2461 TATATTCGAGTTTGTGGTCTG 2484  
Db 2801 TATATTCGAGTTTGTGGTCTG 2824  
RESULT 2  
ABV25313  
ID ABV25313 standard; cDNA; 4031 BP.  
XX AC ABV25313;  
XX 16-SEP-2002 (first entry)  
XX Human prostate expression marker cDNA 25304.  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX pharmacogenomic marker; gene; ss.  
XX Homo sapiens.  
XX WO200160860-A2.  
XX 23-AUG-2001.  
XX 20-FEB-2001; 2001WO-US051171.  
XX 17-FEB-2000; 2000US-183319P.  
XX 16-MAR-2000; 2000US-189862P.  
XX 25-MAY-2000; 2000US-207454P.  
XX 09-JUN-2000; 2000US-211314P.  
XX 18-JUL-2000; 2000US-219007P.  
XX 13-DEC-2000; 2000US-255281P.  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX Schlegel R, Endege WO, Monahan JE;  
XX WPI; 2001-662795/76.  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
XX prostate cells and correlating with presence of prostate cancer, useful  
XX for detecting presence of prostate cancer, stage of prostate cancer -  
XX Claim 1; Page 4980-4981; 11750pp; English.  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
XX specification or its complement. (I) is useful for:  
XX (a) assessing whether a patient is afflicted with prostate cancer;  
XX (b) monitoring the progression of prostate cancer in a patient;  
XX (c) assessing the efficacy of a test compound to inhibit prostate  
XX cancer in a patient;  
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
XX in a patient;  
XX (e) selecting a composition for inhibiting prostate cancer in a patient;  
XX (f) assessing the prostate cell carcinogenic potential of a compound;  
XX (g) determining whether prostate cancer has metastasized in a patient;  
XX (h) assessing the aggressiveness or indolence of prostate cancer in a  
XX patient;  
XX (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX Sequence 4031-BP; 1105 A; 882 C; 918 G; 1120 T; 6 other;  
Query Match 100.0%; Score 2484; DB 23; Length 4031;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGATGAATCTGCACCTGACCCCTGGTACATAAGATGTTTCTTATCTGCCACATTCATCA 60

|||||  
341 ATGGATGAATCTGCACTGACCCCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 400  
QY  
61 GAATACAGTGTGGTGGATGAAGACACAAAGTGAAGATGGGGTGAAGTGGCTTTAGA 120  
Db  
401 GAATACAGTGTGGTGGATGAAGACACAAAGTGAAGATGGGGTGAAGTGGCTTTAGA 460  
QY  
121 CCCACCGTCTTCAGATCTGCAACTTTTAAATGGAAGAAAGCCTTAATAGTTCGGAAGAG 180  
Db  
461 CCCACCGTCTTCAGATCTGCAACTTTTAAATGGAAGAAAGCCTTAATAGTTCGGAAGAG 520  
QY  
181 CCATTTGTTGGAAGATGTTGTTACTCCTGCACTCCCAGAGCTGGGACAAATTTTCAAC 240  
Db  
521 CCATTTGTTGGAAGATGTTGTTACTCCTGCACTCCCAGAGCTGGGACAAATTTTCAAC 580  
QY  
241 CCAGTATCCCGTCTTTGGGTTGGGGAATGTTATTTATATCAATGAAGTCAACAAGA 300  
Db  
581 CCAGTATCCCGTCTTTGGGTTGGGGAATGTTATTTATATCAATGAAGTCAACAAGA 640  
QY  
301 CACCGGGATGGCTTGCAAGACGCCCTTTCTTACGTTCTTTTATCAAGAGCGAGATGTG 360  
Db  
641 CACCGGGATGGCTTGCAAGACGCCCTTTCTTACGTTCTTTTATCAAGAGCGAGATGTG 700  
QY  
361 CATAAGGGCATGTTTGCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 420  
Db  
701 CATAAGGGCATGTTTGCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 760  
QY  
421 GAGGCAATTCAGAAAGTGGCTGCTGAATTTAAACCCCTGATGGTCTGCCCGAGCGAATCA 480  
Db  
761 GAGGCAATTCAGAAAGTGGCTGCTGAATTTAAACCCCTGATGGTCTGCCCGAGCGAATCA 820  
QY  
481 AAGCGCTTTAAAGTGAAGAAAGAAAGCTTAAAGGATCTTCAAGAAATGGTTGCCACT 540  
Db  
821 AAGCGCTTTAAAGTGAAGAAAGAAAGCTTAAAGGATCTTCAAGAAATGGTTGCCACT 880  
QY  
541 GTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTAAACCTGTTCAACAGCTTC 600  
Db  
881 GTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTAAACCTGTTCAACAGCTTC 940  
QY  
601 TTTTGGAAATTCAAATTCAAAGGTCAAACCTTGAGATGGTTAAAGTGAACGTGAGAG 660  
Db  
941 TTTTGGAAATTCAAATTCAAAGGTCAAACCTTGAGATGGTTAAAGTGAACGTGAGAG 1000  
QY  
661 AATTTGCGGCTCTGTTTCTACAGTTTCATAGATCCCATATTCAGTATCTGCTGCTCACT 720  
Db  
1001 AATTTGCGGCTCTGTTTCTACAGTTTCATAGATCCCATATTCAGTATCTGCTGCTCACT 1060  
QY  
721 TTCAATCTCTTCTGCCATTAACATCAAGCAACCATACATTTGCTTCAGGCAATATCTCAAC 780  
Db  
1061 TTCAATCTCTTCTGCCATTAACATCAAGCAACCATACATTTGCTTCAGGCAATATCTCAAC 1120  
QY  
781 ATCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTTCATACGACGAAGCTC 840  
Db  
1121 ATCCCAATCTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTTCATACGACGAAGCTC 1180  
QY  
841 GATGAACACCCAGATGGAGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db  
1181 GATGAACACCCAGATGGAGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1240  
QY  
901 GTTGAATTTACTCGACAGCAGCAATTTCTGGAGATCTTCTGGAAGGCAACAGCTTCTAGG 960  
Db  
1241 GTTGAATTTACTCGACAGCAGCAATTTCTGGAGATCTTCTGGAAGGCAACAGCTTCTAGG 1300  
QY  
961 AGTGGAAAAACCTTGTGCTCGGCGAGGACTTTTGTGAGTGTGGTAGATCTGTGCT 1020  
Db  
1301 AGTGGAAAAACCTTGTGCTCGGCGAGGACTTTTGTGAGTGTGGTAGATCTGTGCT 1360  
QY  
1021 ACCAATGTATCCAGACATCTTTGATTAACCTGTTGGAATCTCTATGATCCGATATC 1080  
Db  
1361 ACCAATGTATCCAGACATCTTTGATTAACCTGTTGGAATCTCTATGATCCGATATC 1420  
QY  
1081 GAAAGTCACTTCAATGGTGAACAACTGGGCAACCTTAAGAAAGATGAGAGCCTGTGGAGT 1140  
|||||

Db  
1421 GAAGTCACTCAATGCTGAACAACTGGGCAAACTTAAGAGAATGAGAGCCTGTGGAGT 1480  
QY  
1141 GTAGCAAGAGTGTATTAGTAATGTACGAAAACTATGGTGTGTCGAGTGGATTTT 1200  
Db  
1481 GTAGCAAGAGTGTATTAGTAATGTACGAAAACTATGGTGTGTCGAGTGGATTTT 1540  
QY  
1201 GCACAGCCATTTTCTTAAAGGAATATTATTAGAAAGCCAAAGTCAGAAACCGGTGTCTGT 1260  
Db  
1541 GCACAGCCATTTTCTTAAAGGAATATTATTAGAAAGCCAAAGTCAGAAACCGGTGTCTGT 1600  
QY  
1261 CTACTTTCCCTGGAGCAAGCGTTGTTA CCAGCTATATCTTCTTCAAGACCCAGTAGTGT 1320  
Db  
1601 CTACTTTCCCTGGAGCAAGCGTTGTTA CCAGCTATATCTTCTTCAAGACCCAGTAGTGT 1660  
QY  
1321 GCTGATGAAGTAGAGACACAGCTCCATTAATAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db  
1661 GCTGATGAAGTAGAGACACAGCTCCATTAATAGTCCAGAAATGCAACAGATGAATCCCTA 1720  
QY  
1381 CGAAGGAGGTTGATTGCAAAATCTGGCTGAGCATAATTCTATTCACTGCTAGCAAGTCTCTGT 1440  
Db  
1721 CGAAGGAGGTTGATTGCAAAATCTGGCTGAGCATAATTCTATTCACTGCTAGCAAGTCTCTGT 1780  
QY  
1441 GCCATTTATGTCACACACATTTGCTGCTGCTGCTCTCTA CAGACAGGCGAGGAAT 1500  
Db  
1781 GCCATTTATGTCACACACATTTGCTGCTGCTGCTCTCTA CAGACAGGCGAGGAAT 1840  
QY  
1501 GATCTCTCACATTTGGTTCGAAGACTTCTTTGTGATGAAGAGGAGTCTGCTGCTGAT 1560  
Db  
1841 GATCTCTCACATTTGGTTCGAAGACTTCTTTGTGATGAAGAGGAGTCTGCTGCTGAT 1900  
QY  
1561 TTTTGACCTGGGGTCTTCAGGAAATTCAGAAAGTGTAGTAATGCATGCCTGCCATACAGCTGCTG 1620  
Db  
1901 TTTTGACCTGGGGTCTTCAGGAAATTCAGAAAGTGTAGTAATGCATGCCTGCCATACAGCTGCTG 1960  
QY  
1621 GGAATTTGTCACAATCA CCCCACACTAGCAGGAAAGTGTATTTTATACCCCCAGC 1680  
Db  
1961 GGAATTTGTCACAATCA CCCCACACTAGCAGGAAAGTGTATTTTATACCCCCAGC 2020  
QY  
1681 ACAACTGCTCCATCAGTCTTCGAACTCAACTTTACGCAATGGGTACTTCATGCTTTT 1740  
Db  
2021 ACAACTGCTCCATCAGTCTTCGAACTCAACTTTACGCAATGGGTACTTCATGCTTTT 2080  
QY  
1741 ATCATGGAGGCCCATCATAGCTTTGCGACCTTTTATGCGAGTCTGAAACAGAGGGAGTCTGGGG 1800  
Db  
2081 ATCATGGAGGCCCATCATAGCTTTGCGACCTTTTATGCGAGTCTGAAACAGAGGGAGTCTGGGG 2140  
QY  
1801 GGTCCCACTAGCACCCCACTAACTGATCAGCAGGAGCAGTGTGCGGAAGGGCGCC 1860  
Db  
2141 GGTCCCACTAGCACCCCACTAACTGATCAGCAGGAGCAGTGTGCGGAAGGGCGCC 2200  
QY  
1861 AGCCTGTGTACTTCTCTCCAAATGAGGCAACATCTCACTGCTTCCAGACATTTTAC 1920  
Db  
2201 AGCCTGTGTACTTCTCTCCAAATGAGGCAACATCTCACTGCTTCCAGACATTTTAC 2260  
QY  
1921 CAAAGTCTGCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATCTTACAGTGGCAGAG 1980  
Db  
2261 CAAAGTCTGCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATCTTACAGTGGCAGAG 2320  
QY  
1981 CACGATGACCGAAGATATCAGTCTTGTGTGAGCAGCAGTGTGGAACAGAAAGCTT 2040  
Db  
2321 CACGATGACCGAAGATATCAGTCTTGTGTGAGCAGCAGTGTGGAACAGAAAGCTT 2380  
QY  
2041 CCAGAACCTTTTCTTTGGAGAGTGAAGAGATGAAGACAGTACATCTTGGGAGGAA 2100  
Db  
2381 CCAGAACCTTTTCTTTGGAGAGTGAAGAGATGAAGACAGTACATCTTGGGAGGAA 2440  
QY  
2101 CAGCGAGATGTCTACTGTAAGGTGAGCCAAATCCAAAGGAGCACAGCAGTTTATCACCTTC 2160  
Db  
2441 CAGCGAGATGTCTACTGTAAGGTGAGCCAAATCCAAAGGAGCACAGCAGTTTATCACCTTC 2500  
QY  
2161 TTACAGAGACTCCTTTGGGCTTTTGTGTGAGGCTTACAGCTGTGCTGCCATCTTTGTTTAC 2220  
Db  
2501 TTACAGAGACTCCTTTGGGCTTTTGTGTGAGGCTTACAGCTGTGCTGCCATCTTTGTTTAC 2560

QY 2221 AACTTCAGTGGCTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTAATA 2280  
DB 2561 AACTTCAGTGGCTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTAATA 2620  
QY 2281 ACCGAAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCCACATATGCTTGTGAAG 2340  
DB 2621 ACCGAAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCCACATATGCTTGTGAAG 2680  
QY 2341 AATGCTGTCAAAATGTTTAAAGATATGCGGGTCTTCAAGGAGACCAACAAAGAGAGTG 2400  
DB 2681 AATGCTGTCAAAATGTTTAAAGATATGCGGGTCTTCAAGGAGACCAACAAAGAGAGTG 2740  
QY 2401 TCTGTTTGAAGTCTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAATCTTAGAA 2460  
DB 2741 TCTGTTTGAAGTCTGAGCAGCACTTTTCTACCTCAATGCAACCGACAAATCTTAGAA 2800  
QY 2461 TATATCTGAGTTTGTGGTGCTG 2484  
DB 2801 TATATCTGAGTTTGTGGTGCTG 2824

RESULT 3

ID ABK94821 standard; DNA; 2487 BP.

AC ABK94821;

XX 29-AUG-2002 (first entry)

DE DNA encoding human glycerol-3-phosphate acyltransferase hGPAT.

XX Fatty acid regulated gene; polyunsaturated fatty acid disorder;

KW PUPA disorder; eczema; cardiovascular disorder; hypertriglyceridaemia;

KW dyslipidaemia; atherosclerosis; coronary artery disease; inflammation;

KW cerebrovascular disease; peripheral vascular disease; rheumatoid arthritis;

KW sinusitis; aschma; pancreatitis; osteoarthritis; cachexia; anorexia;

KW acne; body weight disorder; obesity; cachexia; anorexia;

KW psychiatric disorder; cancer; cystic fibrosis; pre-menstrual syndrome;

KW diabetes; diabetic complication; genetic polymorphism; gene; ds.

XX Homo sapiens.

OS WO200240666-A2.

PN 23-MAY-2002.

XX 19-NOV-2001; 2001WO-CA01632.

XX 17-NOV-2000; 2000US-248589P.

XX (XENO-) XENON GENETICS INC.

XX Winther MD, Goldberg YP, Knickle LC, Haardt M, Allen SJ, Ponton A;

PI De Antueno RJ, Jenkins DK, Nwaka SO;

XX WPI; 2002-508327/54.

XX P-PSDB; ABG66665.

XX Novel isolated polypeptide segment encoded by fat regulated genes,

PT useful for diagnosing the presence of or a predisposition for a

PT disorder involving fatty acid regulated genes in a subject -

XX Claim 1; Fig 28; 225pp; English.

XX The invention describes an isolated polypeptide segment (I) whose genes

CC are fat regulated. (I) or the polynucleotide encoding it (II) are useful

CC for diagnosing the presence of or a predisposition for a disorder

CC involving fatty acid regulated genes in a subject. A composition

CC containing (I) or (II) is useful for treating a disorder involving fatty

CC acid regulated genes, where the disorder is selected from a

CC polyunsaturated fatty acid (PUFA) disorder, eczema, cardiovascular

CC disorders (such as hypertriglyceridaemia, dyslipidaemia, atherosclerosis,

CC coronary artery disease, cerebrovascular disease or peripheral vascular

CC disease), inflammation (such as sinusitis, asthma, pancreatitis,

CC osteoarthritis, rheumatoid arthritis or acne), body weight disorders

CC (such as obesity, cachexia or anorexia), psychiatric disorders, cancer,

CC cystic fibrosis, pre-menstrual syndrome, diabetes, and diabetic

CC complications. (I) or (II) is useful as research agent and materials for

CC discovery of treatments and diagnostics for a disease, particularly human

CC disease. (II) is useful for constructing nucleotide probes and primers,

CC for detecting genetic polymorphism, for detecting changes in the level of

CC expression of (II), and as a diagnostic tool. This sequence encodes

CC a protein regulated by fatty acids.

XX

QY Sequence 2487 BP; 706 A; 548 C; 570 G; 663 T; 0 other;

Best Local Similarity 99.9%; Score 2482.4; DB 24; Length 2487;

Matches 2483; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGATGAATCTGCACCTGACCTTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCA 60

DB 1 ATGGATGAATCTGCACCTGACCTTTGGTACAAATAGATGTTCTTATCTGCCACATTCATCA 60

QY 61 GAATACAGTGTGTCGATGTAAGCACAACAGTGGGATGGGTGAGTGGCTTTAGA 120

DB 61 GAATACAGTGTGTCGATGTAAGCACAACAGTGGGATGGGTGAGTGGCTTTAGA 120

QY 121 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCTTAATGAGTCGGAAGAGG 180

DB 121 CCCACCGTCTTCAGATCTGCAACTTTAAATGGAAGAAAGCCTTAATGAGTCGGAAGAGG 180

QY 181 CCATTTGTTGAAGATGTTGTTTACTCTGCACTCCCGAGAGCTGGGACAAATTTTCAAC 240

DB 181 CCATTTGTTGAAGATGTTGTTTACTCTGCACTCCCGAGAGCTGGGACAAATTTTCAAC 240

QY 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTTATATCAATGAACATCAACAAGA 300

DB 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTTATATCAATGAACATCAACAAGA 300

QY 301 CACCGCGATGGCTTGAAGAGCGCTTCTTACCTCTTTTATTTTCAAGCGGAGATGTG 360

DB 301 CACCGCGATGGCTTGAAGAGCGCTTCTTACCTCTTTTATTTTCAAGCGGAGATGTG 360

QY 361 CATAGGGGCAATGTTTGGCCACCAATGTGCTGAAATGTGCTGAACAGCAGTAGAGTACAA 420

DB 361 CATAGGGGCAATGTTTGGCCACCAATGTGCTGAAATGTGCTGAACAGCAGTAGAGTACAA 420

QY 421 GAGGCAATTCAGAAAGTGGCTGCTGAATTAACCCCTGATGTTCTGCCAGCAGCAATCA 480

DB 421 GAGGCAATTCAGAAAGTGGCTGCTGAATTAACCCCTGATGTTCTGCCAGCAGCAATCA 480

QY 481 AAAGCGGTTTAAAGTGAAGAAAGCTTAAAGGATTTCTCAAGAAATGGTTGCCACT 540

DB 481 AAAGCGGTTTAAAGTGAAGAAAGCTTAAAGGATTTCTCAAGAAATGGTTGCCACT 540

QY 541 GTCTCACCGGCAATGATCAGACTGACCTGGGTTGGGTTGCTGCTTCAAGAAATGGTTGCCACT 600

DB 541 GTCTCACCGGCAATGATCAGACTGACCTGGGTTGGGTTGCTGCTTCAAGAAATGGTTGCCACT 600

QY 601 TTTTGGAAATTCATAATTCACAAAGGTCACCTGAGATGGTTTAAAGCTGCAACTGAGACG 660

DB 601 TTTTGGAAATTCATAATTCACAAAGGTCACCTGAGATGGTTTAAAGCTGCAACTGAGACG 660

QY 661 AATTGGCGCTTCTGTTTCTACAGATTCATAGATCCCATATTTGACTATCTGCTGCTCACT 720

DB 661 AATTGGCGCTTCTGTTTCTACAGATTCATAGATCCCATATTTGACTATCTGCTGCTCACT 720

QY 721 TTCAATCTCTTCTGCCATAACATCAAGAGCACCACATATGCTTTCAGGCAATAATCTCAAC 780

DB 721 TTCAATCTCTTCTGCCATAACATCAAGAGCACCACATATGCTTTCAGGCAATAATCTCAAC 780

QY 781 ATCCCAATCTTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTTCATACGACGAGGCTC 840

DB 781 ATCCCAATCTTTCAGTACCTTGATCCATAAGCTTGGGGCTTCTTTCATACGACGAGGCTC 840

```
QY 841 GATGAAACACCATGAGCGGAAGAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900
DB |||||
DB 841 GATGAAACACCATGAGCGGAAGAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900
QY 901 GTTGAATTAATCTCGACAGCAGCAATCTCTGGAGATCTTCCTGGAAGCAGACACCTTCTAG 960
DB |||||
DB 901 GTTGAATTAATCTCGACAGCAGCAATCTCTGGAGATCTTCCTGGAAGCAGACACCTTCTAG 960
QY 961 AGTGGAAACCTCTTGTGCTCGGCAGGACTTTTGTCTAGTTGTGTAGTACTCTGTCT 1020
DB |||||
DB 961 AGTGGAAACCTCTTGTGCTCGGCAGGACTTTTGTCTAGTTGTGTAGTACTCTGTCT 1020
QY 1021 ACCAATGTCTATCCAGACATCTTTGATAATACCTGTTGGAACTCTCTATGATCCATTATC 1080
DB |||||
DB 1021 ACCAATGTCTATCCAGACATCTTTGATAATACCTGTTGGAACTCTCTATGATCCATTATC 1080
QY 1081 GAAGGTCACCTAATGGTGAACAACTGGGCAAACTTAAGAAAGATGAGAGCCTGTGGAGT 1140
DB |||||
DB 1081 GAAGGTCACCTAATGGTGAACAACTGGGCAAACTTAAGAAAGATGAGAGCCTGTGGAGT 1140
QY 1141 GTAGCAAGAGTGTATTAGAAATGTTACGNAAAAACTATGGTGTGTCTCGAGTGGATTTT 1200
DB |||||
DB 1141 GTAGCAAGAGTGTATTAGAAATGTTACGNAAAAACTATGGTGTGTCTCGAGTGGATTTT 1200
QY 1201 GCACAGCATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCT 1260
DB |||||
DB 1201 GCACAGCATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGTCTGCT 1260
QY 1261 CTACTTTCCCTGGAGCAAGCGTTGTTACCAAGCTATATCTTCAAGACCCAGTGATGCT 1320
DB |||||
DB 1261 CTACTTTCCCTGGAGCAAGCGTTGTTACCAAGCTATATCTTCAAGACCCAGTGATGCT 1320
QY 1321 GCTGATGAAGGTAGACACAGCTCCATTATGATCCAGAAATCCACAGATGAATCCCTA 1380
DB |||||
DB 1321 GCTGATGAAGGTAGACACAGCTCCATTATGATCCAGAAATCCACAGATGAATCCCTA 1380
QY 1381 CGAAGGAGGTTGATGTCAAATCTGGCTGAGCATATTTCTATTCTACTGCTAGCAAGTCTGT 1440
DB |||||
DB 1381 CGAAGGAGGTTGATGTCAAATCTGGCTGAGCATATTTCTATTCTACTGCTAGCAAGTCTGT 1440
QY 1441 GCATTATGTCACACACATGTTGGCTGCTGCTCTCTCTACAGACACAGGCGGGAAT 1500
DB |||||
DB 1441 GCATTATGTCACACACATGTTGGCTGCTGCTCTCTCTCTACAGACACAGGCGGGAAT 1500
QY 1501 GATCTCTCCACATGTCGAGACTTCTTTGTGATGAAGAGAGAGTCTCTGGCTCTGTAT 1560
DB |||||
DB 1501 GATCTCTCCACATGTCGAGACTTCTTTGTGATGAAGAGAGAGTCTCTGGCTCTGTAT 1560
QY 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAAGTGTAGTATGATGCTATGCCATACAGCTGCTG 1620
DB |||||
DB 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAAGTGTAGTATGATGCTATGCCATACAGCTGCTG 1620
QY 1621 GGAAATTTGTCTCAATACACCCACACTAGCAGGAAGCATGAGTTTTTTATCACCCCCAGC 1680
DB |||||
DB 1621 GGAAATTTGTCTCAATACACCCACACTAGCAGGAAGCATGAGTTTTTTATCACCCCCAGC 1680
QY 1681 ACACTGTCCCATCAGTCTTGGAACTCACTTCTACAGCAATGGGGTACTTATGTCCTTT 1740
DB |||||
DB 1681 ACACTGTCCCATCAGTCTTGGAACTCACTTCTACAGCAATGGGGTACTTATGTCCTTT 1740
QY 1741 ATCATGGAGCCATCATAGCTTTCAGCTTTATGTCAGTTCTGAACAGAGGGGACTGGGG 1800
DB |||||
DB 1741 ATCATGGAGCCATCATAGCTTTCAGCTTTATGTCAGTTCTGAACAGAGGGGACTGGGG 1800
QY 1801 GGTCCCATAGCACCCACCTAACCTGATCAGCAGGAGCAGCTGGTGGGAAGGGCGGC 1860
DB |||||
DB 1801 GGTCCCATAGCACCCACCTAACCTGATCAGCAGGAGCAGCTGGTGGGAAGGGCGGC 1860
QY 1861 AGCCTGTGCTACCTTCTCTCAATGAAGCAGCAGCTCTACTGCTTGGCCAGCATTTTAC 1920
DB |||||
DB 1861 AGCCTGTGCTACCTTCTCTCAATGAAGCAGCAGCTCTACTGCTTGGCCAGCATTTTAC 1920
```

```
QY 1921 CAAAGTCGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTTACAGTGGCAGAG 1980
DB |||||
DB 1921 CAAAGTCGCCATGAAACAGTAGGAAAGTTTATCCAGTATGGCATTTTACAGTGGCAGAG 1980
QY 1981 CACGATGACCAGGAAGATATCAGTCTTAGTCTTGTGACGACAGTGGGACAAAGAGCTT 2040
DB |||||
DB 1981 CACGATGACCAGGAAGATATCAGTCTTAGTCTTGTGACGACAGTGGGACAAAGAGCTT 2040
QY 2041 CCAGAACCTTTCTCTTGGAGAAAGTATGATGAAGAAAGATGAAGACAGTACTTTGGGGAGGAA 2100
DB |||||
DB 2041 CCAGAACCTTTCTCTTGGAGAAAGTATGATGAAGAAAGATGAAGACAGTACTTTGGGGAGGAA 2100
QY 2101 CAGCGAGATTGCTACTCTGAAGGTGAGCCCAATCCAAAGGAGCACCAGCAGTTTATCACCTTC 2160
DB |||||
DB 2101 CAGCGAGATTGCTACTCTGAAGGTGAGCCCAATCCAAAGGAGCACCAGCAGTTTATCACCTTC 2160
QY 2161 TTACAGAGACTCTCTTGGGCTTTTGTGAGGCTTACAGCTCTGCTGCCATCTTTTGTTCAC 2220
DB |||||
DB 2161 TTACAGAGACTCTCTTGGGCTTTTGTGAGGCTTACAGCTCTGCTGCCATCTTTTGTTCAC 2220
QY 2221 AACTTCAGTGTCTCTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTTAATA 2280
DB |||||
DB 2221 AACTTCAGTGTCTCTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTTAATA 2280
QY 2281 ACCAGAAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCCACATATTTGTCTGTGAAG 2340
DB |||||
DB 2281 ACCAGAAACAGAAAGAAATGTTGCAGTATATGCTGAGAGTGCCACATATTTGTCTGTGAAG 2340
QY 2341 AATGCTGTGAAATGTTTAAAGGATATTTGGGGTTTTTCAAGGAGACCCAAACAAAGAGAGTG 2400
DB |||||
DB 2341 AATGCTGTGAAATGTTTAAAGGATATTTGGGGTTTTTCAAGGAGACCCAAACAAAGAGAGTG 2400
QY 2401 TCTGTTTTAGAACTGAGCAGACCTTTTCTACCTCAATGCAACCGACAAAACTTCTAGAA 2460
DB |||||
DB 2401 TCTGTTTTAGAACTGAGCAGACCTTTTCTACCTCAATGCAACCGACAAAACTTCTAGAA 2460
QY 2461 TATATTCTCAGTTTTTGTGTGCTG 2484
DB |||||
DB 2461 TATATTCTCAGTTTTTGTGTGCTG 2484
```

## RESULT 4

AAL55475  
ID AAL55475 standard; DNA; 2487 BP.

XX AAL55475;

XX 22-MAY-2003 (first entry)

XX GPAM related DNA sequence, SEQ ID No 1.

DE Antidiabetic; nephrotropic; neuroprotective; ophthalmological; human;  
DE mitochondrial sn-glycerol-3-phosphate acyltransferase; GPAM;  
KW diabetic complication; retinopathy; neuropathy; enzyme; gene; ds.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH 1..2487  
FT CDS /tag= a  
FT /product= "GPAM-related protein"

XX W02003008590-A1.

XX 30-JAN-2003.

XX 16-JUL-2002; 2002WO-JP07189.

XX 16-JUL-2001; 2001JP-0215337.

XX (KISP ) KISSEI PHARM CO LTD.

XX Sakamoto S, Onota H, Sugano S, Nakamura Y;

XX WPI; 2003-229583/22.  
DR P-PSDB; AAO27073.  
XX  
XX Human mitochondrial sn-glycerol-3-phosphate acyltransferase and  
PT antagonists for treatment and prevention of diabetic complications  
XX  
XX Claim 4; Page 33-40; 56pp; Japanese.  
XX  
XX The invention relates to a novel protein having human mitochondrial sn-  
CC glycerol-3-phosphate acyltransferase (GPAM) activity. The novel protein  
CC with GPAM activity can be used in the prevention and treatment of  
CC diabetic complications, including retinopathy and neuropathy, by  
CC administration of antagonists to human GPAM. This polynucleotide sequence  
CC represents a DNA sequence encoding a protein relating to the GPAM  
XX activity protein of the invention.  
XX  
SQ Sequence 2487 BP; 704 A; 549 C; 571 G; 563 T; 0 other;

Query Match 99.9%; Score 2480.8; DB 25; Length 2487;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2482; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGATGAATCTGCACTGACCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 60  
Db 1 ATGGATGAATCTGCACTGACCTTGGTACAAATAGATGTTTCTTATCTGCCACATTCATCA 60

Qy 61 GAATACAGTGTGGTGCAGTGAAGCACACAAGTGAAGGAATGGGTGAGTGGCTTTAGA 120  
Db 61 GAATACAGTGTGGTGCAGTGAAGCACACAAGTGAAGGAATGGGTGAGTGGCTTTAGA 120

Qy 121 CCCACCGCTTTCCAGATCTGCAACTTTAAATGGAAGAAAGCCTTAATCAGTCGGAAGAG 180  
Db 121 CCCACCGCTTTCCAGATCTGCAACTTTAAATGGAAGAAAGCCTTAATCAGTCGGAAGAG 180

Qy 181 CCATTTGTTGAAGATGTTTACTCTCGACTCCAGAGCTGGGACAAATTTTCAAC 240  
Db 181 CCATTTGTTGAAGATGTTTACTCTCGACTCCAGAGCTGGGACAAATTTTCAAC 240

Qy 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAACCTCACACAAGA 300  
Db 241 CCCAGTATCCCGTCTTTGGGTTTGGGAATGTTATTTATATCAATGAACCTCACACAAGA 300

Qy 301 CACCGGATGGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTG 360  
Db 301 CACCGGATGGCTTGCAGAGCGCTTTCTTACGTTCTTTTATTTCAAGACGAGATGTG 360

Qy 361 CATAGGCGATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 420  
Db 361 CATAGGCGATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 420

Qy 421 GAGGCAATTGCAGAGTGGCTGCTGAATTAACCTGATGTTCTGCCACGACGCAATCA 480  
Db 421 GAGGCAATTGCAGAGTGGCTGCTGAATTAACCTGATGTTCTGCCACGACGCAATCA 480

Qy 481 AAAGCCGTTAAACAAAGTGAAGAAAGCAATTAAGGATTTCTCAAGAAATGGTGGCCACT 540  
Db 481 AAAGCCGTTAAACAAAGTGAAGAAAGCAATTAAGGATTTCTCAAGAAATGGTGGCCACT 540

Qy 541 GTCTCACCGCAATGATCAGACTGCTGGGTGGGTGCTGCTTAAACTGTTTCAACAGCTTC 600  
Db 541 GTCTCACCGCAATGATCAGACTGCTGGGTGGGTGCTGCTTAAACTGTTTCAACAGCTTC 600

Qy 601 TTTTGGACATTCAAATTCAGAAAGTCAACTTGATGATGTTTAAAGTGCAGCTGAGACG 660  
Db 601 TTTTGGACATTCAAATTCAGAAAGTCAACTTGATGATGTTTAAAGTGCAGCTGAGACG 660

Qy 661 AATTGCGCTTCTGTTTCTACCACTTCATAGATCCCATATTTGATATCTGCTCACT 720  
Db 661 AATTGCGCTTCTGTTTCTACCACTTCATAGATCCCATATTTGATATCTGCTCACT 720

Qy 721 TTCAATTCCTTCTGCCATAAATCAAGCACCATACTGCTTTCAGGCAATTAATCTCAAC 780  
Db 721 TTCAATTCCTTCTGCCATAAATCAAGCACCATACTGCTTTCAGGCAATTAATCTCAAC 780

Db 721 TTCAATTCCTTCTGCCATAAATCAAGCACCATACTGCTTTCAGGCAATTAATCTCAAC 780  
Qy 781 ATCCCAATCTTCAGTACCTTGATCATAAGCTTGGGGCTTCTTTCATACAGCAGAGGCTC 840  
Db 781 ATCCCAATCTTCAGTACCTTGATCATAAGCTTGGGGCTTCTTTCATACAGCAGAGGCTC 840

Qy 841 GATGAAACACACAGATGGAACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db 841 GATGAAACACACAGATGGAACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900

Qy 901 GTTGAATTTACTTCCACAGCAGCAATTTCTGGAGATCTTCTCGGAAGGACACAGTCTTAGG 960  
Db 901 GTTGAATTTACTTCCACAGCAGCAATTTCTGGAGATCTTCTCGGAAGGACACAGTCTTAGG 960

Qy 961 AGTGGAAACCTCTTGTGCTCGGCGAGGACTTTTGTGAGTGTGGTAGACTCTGTCT 1020  
Db 961 AGTGGAAACCTCTTGTGCTCGGCGAGGACTTTTGTGAGTGTGGTAGACTCTGTCT 1020

Qy 1021 ACCAATGTCTCCAGACATCTTGTGATTAATACCTGTTTGGAAATCTCTATGATCGCATATC 1080  
Db 1021 ACCAATGTCTCCAGACATCTTGTGATTAATACCTGTTTGGAAATCTCTATGATCGCATATC 1080

Qy 1081 GAAGGTCACTACAATGTTGAACAACTGGGCAACCTTAAGAAAGATGAGAGCTGTGGAGT 1140  
Db 1081 GAAGGTCACTACAATGTTGAACAACTGGGCAACCTTAAGAAAGATGAGAGCTGTGGAGT 1140

Qy 1141 GTAGCAAGAGGTGTTATTAGAACTGTACGAAACAACTATGTTGTTGTTCCGAGTGGATTTT 1200  
Db 1141 GTAGCAAGAGGTGTTACTAGAACTGTACGAAACAACTATGTTGTTGTTCCGAGTGGATTTT 1200

Qy 1201 GCACAGCCATTTTCTTAAAGGAAATTTTAAAGAGCCAAAGTCAAGAACCCGGTGTCTGCT 1260  
Db 1201 GCACAGCCATTTTCTTAAAGGAAATTTTAAAGAGCCAAAGTCAAGAACCCGGTGTCTGCT 1260

Qy 1261 CTACTTTCCCTGGAGAGAGCGTTGTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320  
Db 1261 CTACTTTCCCTGGAGAGAGCGTTGTTACAGCTATATCTTCTTCAAGACCCAGTGTGCT 1320

Qy 1321 GCTGATCAAGGTGAGACACGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db 1321 GCTGATCAAGGTGAGACACGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380

Qy 1381 CGAAGGAGGTGATTTGCAAAATCTGGCTCAGCATATTTCTATTCTAGTCAAGTCTCTGT 1440  
Db 1381 CGAAGGAGGTGATTTGCAAAATCTGGCTCAGCATATTTCTATTCTAGTCAAGTCTCTGT 1440

Qy 1441 GCCATTATGTCACACACATTTGTGGCTTGGCTGCTCTCTACACACACAGCAGGGAAT 1500  
Db 1441 GCCATTATGTCACACACATTTGTGGCTTGGCTGCTCTCTCTACACACACAGCAGGGAAT 1500

Qy 1501 GATCTCTCCATTTGTCGAGAGCTTCTTGTGATGAAGGAGGAGTCCCTGGCTCGTAT 1560  
Db 1501 GATCTCTCCATTTGTCGAGAGCTTCTTGTGATGAAGGAGGAGTCCCTGGCTCGTAT 1560

Qy 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAAGTGTAGTAAATGTCATGTCATACAGTCTGTG 1620  
Db 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAAGTGTAGTAAATGTCATGTCATACAGTCTGTG 1620

Qy 1621 GGAATTTGTGTCAAAATCACCACTAGCAGGAAACGATGATTTTTTATCACCCCCAGC 1680  
Db 1621 GGAATTTGTGTCAAAATCACCACTAGCAGGAAACGATGATTTTTTATCACCCCCAGC 1680

Qy 1681 ACACCTGTCCTTTCAGTCTTTCGAACTCAATTTCTACAGCAATGGGGTACTTCATGTCTTT 1740  
Db 1681 ACACCTGTCCTTTCAGTCTTTCGAACTCAATTTCTACAGCAATGGGGTACTTCATGTCTTT 1740

Qy 1741 ATCATGAGGCGCATCATAGCTTTCAGGCTTTTATGAGTCTTCTGAACAAGAGGGGACTGGGG 1800  
Db 1741 ATCATGAGGCGCATCATAGCTTTCAGGCTTTTATGAGTCTTCTGAACAAGAGGGGACTGGGG 1800

Qy 1801 GGTCCTCACTAGCAGCCCCACCTTAACCTGATCAGCAGGAGCAGCTGGTCGGAAGCGCGCC 1860  
Db 1801 GGTCCTCACTAGCAGCCCCACCTTAACCTGATCAGCAGGAGCAGCTGGTCGGAAGCGCGCC 1860



```
QY 1861 AGCTGTGCTACCTTCTCTCAATGAAGGACCACTCTCACTGCTTGGCCAGACATTTTAC 1920
Db |||||
QY 1861 AGCTGTGCTACCTTCTCTCAATGAAGGACCACTCTCACTGCTTGGCCAGACATTTTAC 1920
Db |||||
QY 1921 CAAGTCTGCCATGAACAGTAGGAAGTTTATCAGATATGGCATTTTACAGTGGCAGAG 1980
Db |||||
QY 1921 CAAGTCTGCCATGAACAGTAGGAAGTTTATCAGATATGGCATTTTACAGTGGCAGAG 1980
QY 1981 CAGGATGACCAAGAGATATCAGTCTCTAGTCTTGTCTGAGCAGAGTGGGACAGAGCTT 2040
Db |||||
QY 1981 CAGGATGACCAAGAGATATCAGTCTCTAGTCTTGTCTGAGCAGAGTGGGACAGAGCTT 2040
QY 2041 CCAGAACCTTTTGTCTTGGAGAGTGATGAAGAGATGAAGAGATGACATTTTGGGGAGGAA 2100
Db |||||
QY 2041 CCTGAACCTTTTGTCTTGGAGAGTGATGAAGAGATGAAGAGATGACATTTTGGGGAGGAA 2100
QY 2101 CAGGAGATTTGCTACCTGAAGGTGAGCCAAATCCAAAGAGCACCAGCAGTTTATCAGCTTC 2160
Db |||||
QY 2101 CAGGAGATTTGCTACCTGAAGGTGAGCCAAATCCAAAGAGCACCAGCAGTTTATCAGCTTC 2160
QY 2161 TTACAGAGACTCTTGGGCTTTTGTCTGAGGCTTACAGCTCTGCTGCCATCTTTGTTCAC 2220
Db |||||
QY 2161 TTACAGAGACTCTTGGGCTTTTGTCTGAGGCTTACAGCTCTGCTGCCATCTTTGTTCAC 2220
QY 2221 AACTTCAGTGTCTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTAATA 2280
Db |||||
QY 2221 AACTTCAGTGTCTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCACAAATACCTAATA 2280
QY 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATTTCTTGTGAAG 2340
Db |||||
QY 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATTTCTTGTGAAG 2340
QY 2341 AATGCTGTGAAGATTTTAAAGATATTTGGGCTTTCAGGAGACCAACAAAGAGAGTG 2400
Db |||||
QY 2341 AATGCTGTGAAGATTTTAAAGATATTTGGGCTTTCAGGAGACCAACAAAGAGAGTG 2400
QY 2401 TCTGTTTTAGAACTGAGCAGACATTTTCTACCTCAATGCAACGACAAACAACTTCTAGAA 2460
Db |||||
QY 2401 TCTGTTTTAGAACTGAGCAGACATTTTCTACCTCAATGCAACGACAAACAACTTCTAGAA 2460
QY 2461 TATATTCTGAGTTTTTGTGTGCTG 2484
Db |||||
QY 2461 TATATTCTGAGTTTTTGTGTGCTG 2484

RESULT 5
AAD35221
AC AAD35221 standard; cDNA; 3210 BP.
AC AAD35221;
DE 25-JUL-2002 (first entry)
XX Human TRNFR-6 cDNA.
XX Human; transferase; developmental disorder; cell proliferative disorder;
XX TRNFR-6; neurological disorder; autoimmune disorder; parasitic infection;
XX inflammatory disorder; endocrine; antiparasitic; immunosuppressive;
XX cytosolic; neurological; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 338..2824
XX FT /*tag= a
XX FT /product= "Human TRNFR-6 protein"
XX PN WO200226950-A2.
XX PD 04-APR-2002.
XX FT
XX PF 28-SEP-2001; 2001WO-US30424.
```

```
XX 29-SEP-2000; 2000US-236523P.
PR 06-OCT-2000; 2000US-238481P.
PR 27-OCT-2000; 2000US-244025P.
PR 03-NOV-2000; 2000US-246001P.
PR 09-NOV-2000; 2000US-247931P.
PR 16-NOV-2000; 2000US-249639P.
PR 21-NOV-2000; 2000US-252819P.
XX (INCY-) INCYTE GENOMICS INC.
XX Lal PG, Tang YT, Yue H, Burford N, Gandhi AR, Warren BA, Yao MG;
PI Tribouley CM, Baughn MR, Lee EA, Hafalia AJA, Lu Y, Griffin JA;
PI Sanjanwala MS, Ding L;
XX WPI: 2002-362492/39.
DR P-PSDB; AAE22144.
XX Novel human transferase polypeptides and polynucleotides, useful in
PT treating e.g., cell proliferative and autoimmune disorders -
XX Claim 81; Page 154-155; 168pp; English.
XX The present invention relates to novel human transferases (TRNFR) and
CC polynucleotides encoding such proteins. The TRNFR proteins are useful
CC for treating disorders associated with a decreased expression of
CC functional TRNFR, e.g., cell proliferative, developmental, neurological,
CC autoimmune/inflammatory disorders and parasitic infections. Antagonists
CC of TRNFR proteins are useful in treating disorders associated with
CC increased activity of TRNFR. The present sequence is human TRNFR-6 cDNA.
XX SQ Sequence 3210 BP; 896 A; 721 C; 735 G; 858 T; 0 other;
```

Query Match 99.9%; Score 2480.8; DB 24; Length 3210;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2482; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGGATGAATCTGCACTGACCTTGGTACAATAGATGTTCTTATCTGCCACATTCATCA 60
Db 338 ATGGATGAATCTGCACTGACCTTGGTACAATAGATGTTCTTATCTGCCACATTCATCA 397
QY 61 GAATACAGTGTGGTGCATGTAAGCACACAGTGAAGAAATGGGTGAGTGTGGCTTTAGA 120
Db 398 GAATACAGTGTGGTGCATGTAAGCACACAGTGAAGAAATGGGTGAGTGTGGCTTTAGA 457
QY 121 CCCACCGTCTTCAGATCTGCAACTTTAAATCGAAGAAAGCTAATGAGTCGGAAGG 180
Db 458 CCCACCATCTTCAGATCTGCAACTTTAAATCGAAGAAAGCTAATGAGTCGGAAGG 517
QY 181 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGCAGAGCTGGGACAAATTTTCAAC 240
Db 518 CCATTTGTTGGAAGATGTTGTTACTCTGCACTCCCGCAGAGCTGGGACAAATTTTCAAC 577
QY 241 CCCAGTATCCCGTCTTTGGGTTTGGGAAATGTTATTTATATCAATGAAGAACTCACACA 300
Db 578 CCCAGTATCCCGTCTTTGGGTTTGGGAAATGTTATTTATATCAATGAAGAACTCACACA 637
QY 301 CACCGCGGATGCTGTCAGAGCCCTTTCTTACGTTCTTTTATTTCAGAGCGAGATGTG 360
Db 638 CACCGCGGATGCTGTCAGAGCCCTTTCTTACGTTCTTTTATTTCAGAGCGAGATGTG 697
QY 361 CATAGGGCATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 420
Db 698 CATAGGGCATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAGTACAA 757
QY 421 GAGGCAATTCAGAAAGTGGCTCTGAATTAACCCCTGATGTTCTGCCAGCAGCAATCA 480
Db 758 GAGGCAATTCAGAAAGTGGCTCTGAATTAACCCCTGATGTTCTGCCAGCAGCAATCA 817
QY 481 AAAGCCGTTAACCAAGTGAAGAAAGAAAGCTAAAGGATTTCTTCNAGAAATGGTGCAC 540
Db 818 AAAGCCGTTAACCAAGTGAAGAAAGAAAGCTAAAGGATTTCTTCNAGAAATGGTGCAC 877
```

QY 541 GTCTCACCGGCAATGATCAGACTGACTGGTGGTGGTCTGCTAAACTGTTCAACAGCTTC 600  
Db 878 GTCTCACCGGCAATGATCAGACTGACTGGTGGTGGTCTGCTAAACTGTTCAACAGCTTC 937  
QY 601 TTTTGGAACTTCAAAATTCACAAAGGTCAACTTGGAGATGGTTAAAGCTGCAACTGAGACG 660  
Db 938 TTTTGGAACTTCAAAATTCACAAAGGTCAACTTGGAGATGGTTAAAGCTGCAACTGAGACG 997  
QY 661 AATTGGCCGCTTCTGTTTCTACAGTTCATAGATCCCATATTTGACTATCTGCTGCTCACT 720  
Db 998 AATTGGCCGCTTCTGTTTCTACAGTTCATAGATCCCATATTTGACTATCTGCTGCTCACT 1057  
QY 721 TTCACTCTCTTCTGTCATTAACATCAAGACCATATGCTTCAAGCAATAATCTCAAC 780  
Db 1058 TTCACTCTCTTCTGTCATTAACATCAAGACCATATGCTTCAAGCAATAATCTCAAC 1117  
QY 781 ATCCCAATCTTCAGTACCTTGTATCATAGCTTGGGGCTTCTTCATACGACGAGGCTC 840  
Db 1118 ATCCCAATCTTCAGTACCTTGTATCATAGCTTGGGGCTTCTTCATACGACGAGGCTC 1177  
QY 841 GATGAACACCCAGATGCGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db 1178 GATGAACACCCAGATGCGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1237  
QY 901 GTTGAATTTCTCGACAGCAATTTCTTGAGATCTTCTGGAAGGCACACGTTCTAGG 960  
Db 1238 GTTGAATTTCTCGACAGCAATTTCTTGAGATCTTCTGGAAGGCACACGTTCTAGG 1297  
QY 961 AGTGGAAAACTCTTGTGCTCGGCAGGACTTTTGTCAAGTTGTGGTAGACTCTGTCT 1020  
Db 1298 AGTGGAAAACTCTTGTGCTCGGCAGGACTTTTGTCAAGTTGTGGTAGACTCTGTCT 1357  
QY 1021 ACCAATGTCTATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATCGCATATC 1080  
Db 1358 ACCAATGTCTATCCAGACATCTTGATAATACCTGTTGGAATCTCTATGATCGCATATC 1417  
QY 1081 GAAGTCTACTAATGTTGAACAACTGGGCAAACTTAAGAAATGAGAGCTTGTGGAGT 1140  
Db 1418 GAAGTCTACTAATGTTGAACAACTGGGCAAACTTAAGAAATGAGAGCTTGTGGAGT 1477  
QY 1141 GTACGAGAGTGTATTAGATGTTAGAAATGTTAGAAATGTTAGTGTGTCGAGTGGATTTT 1200  
Db 1478 GTAGCAAGAGTGTATTAGATGTTAGAAATGTTAGAAATGTTAGTGTGTCGAGTGGATTTT 1537  
QY 1201 GCACAGCCATTTCTTAAAGGAATTTTAGAAGCCAAAGTCAGAAACCCGTTGTGCT 1260  
Db 1538 GCACAGCCATTTCTTAAAGGAATTTTAGAAGCCAAAGTCAGAAACCCGTTGTGCT 1597  
QY 1261 CTACTTTCCTGGAGCAAGCGTTGTTACAGCTATCTTCTTCAAGACCCAGTGTGCT 1320  
Db 1598 CTACTTTCCTGGAGCAAGCGTTGTTACAGCTATCTTCTTCAAGACCCAGTGTGCT 1657  
QY 1321 GCTGATGAAGTAGAGACAGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db 1658 GCTGATGAAGTAGAGACAGCTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1717  
QY 1381 CGAAGGAGTGTGATTGCAAAATCTGCTCAGCATATTTCTTCACTGTAGCAAGTCTGCT 1440  
Db 1718 CGAAGGAGTGTGATTGCAAAATCTGCTCAGCATATTTCTTCACTGTAGCAAGTCTGCT 1777  
QY 1441 GCCATTATGTCACACACATTTGTGGCTTGCCTCTCTACAGACACAGCAGGGAATT 1500  
Db 1778 GCCATTATGTCACACACATTTGTGGCTTGCCTCTCTCTACAGACACAGCAGGGAATT 1837  
QY 1501 GATCTCTCCATTTGGTGGAGACTTTCTTTGTGATGAAAGGAGTCTCTGGCTCGTGAT 1560  
Db 1838 GATCTCTCCATTTGGTGGAGACTTTCTTTGTGATGAAAGGAGTCTCTGGCTCGTGAT 1897  
QY 1561 TTTGACCTGGGCTTCTCAGGAAATTCAGAGATGTAGTAATGCTATGCTCATACAGCTGCTG 1620  
Db 1898 TTTGACCTGGGCTTCTCAGGAAATTCAGAGATGTAGTAATGCTATGCTCATACAGCTGCTG 1957  
QY 1621 GGAAATTTGTGTCACAATCACCCACACTAGCAGGAACGATGAGTTTTTTTATCACCCCCAGC 1680

Db 1958 GGAAATTTGTGTCACAATCACCCACACTAGCAGGAACGATGAGTTTTTTTATCACCCCCAGC 2017  
QY 1681 ACAACTGTGCCATCAGTCTTCCGAACTCAACTTCTACAGCAATGGGGTACTTCTATGCTCTTT 1740  
Db 2018 ACAACTGTGCCATCAGTCTTCCGAACTCAACTTCTACAGCAATGGGGTACTTCTATGCTCTTT 2077  
QY 1741 ATCATGGAGGCAATCATAGCTTGCAGCTTTTATGCAGTCTTCTGAAACAGAGGGGACTGGGG 1800  
Db 2078 ATCATGGAGGCAATCATAGCTTGCAGCTTTTATGCAGTCTTCTGAAACAGAGGGGACTGGGG 2137  
QY 1801 GGTCCCACTAGCACCCACCTAATCCTGATCAGCAGAGGAGCTGGTGGGAGGCGGCC 1860  
Db 2138 GGTCCCACTAGCACCCACCTAATCCTGATCAGCAGGAGGAGCTGGTGGGAGGCGGCC 2197  
QY 1861 AGCCTGTGCTACCTTCTCTCAATGAAGGCACTCTCAGTCCCTTCCAGCCCTGCGACATTTTTAC 1920  
Db 2198 AGCCTGTGCTACCTTCTCTCAATGAAGGCACTCTCAGTCCCTTCCAGCCCTGCGACATTTTTAC 2257  
QY 1921 CAACTCTGCCATGAAACAGTAGGAAAGTTTTATCCAGTATGGCATTTCTTACAGTGGCAGAG 1980  
Db 2258 CAACTCTGCCATGAAACAGTAGGAAAGTTTTATCCAGTATGGCATTTCTTACAGTGGCAGAG 2317  
QY 1981 CACGATGACGAGGAAGATATCAGTCTTCTGCTGAGCAGAGTGGGAACAAGAGCTT 2040  
Db 2318 CACGATGACGAGGAAGATATCAGTCTTCTGCTGAGCAGAGTGGGAACAAGAGCTT 2377  
QY 2041 CCAGAACTTTGTCTTGGAGAGTGAAGAACTGAAGACAGTGAATTTGGGGAGGAA 2100  
Db 2378 CCTGAACCTTTGTCTTGGAGAGTGAAGAACTGAAGACAGTGAATTTGGGGAGGAA 2437  
QY 2101 CAGCGAGATTGCTACCTGAAAGGTGAGCAATCCAAAGGAGCAGCAGCTTTTATCACCTTC 2160  
Db 2438 CAGCGAGATTGCTACCTGAAAGGTGAGCAATCCAAAGGAGCAGCAGCTTTTATCACCTTC 2497  
QY 2161 TTACAGAGCTCTTGGGCCCTTTGCTGAGGCCCTTACAGCTCTGCTGCCATCTTTGTTTCAAC 2220  
Db 2498 TTACAGAGCTCTTGGGCCCTTTGCTGAGGCCCTTACAGCTCTGCTGCCATCTTTGTTTCAAC 2557  
QY 2221 AACTTCAGTGGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAATA 2280  
Db 2558 AACTTCAGTGGTCTGTTCCAGAACCTGAGTATCTGCAAAAGTTGCAAAATACCTTAATA 2617  
QY 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTCAGAGTGCACATATGTTGTTGTGAAG 2340  
Db 2618 ACCAGAACAGAAAGAAATGTTGAGTATATGCTCAGAGTGCACATATGTTGTTGTGAAG 2677  
QY 2341 AATGCTGTGAAATGTTTAAAGGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTG 2400  
Db 2678 AATGCTGTGAAATGTTTAAAGGATATTTGGGGTTTTCAAGGAGACCAAAACAAAGAGAGTG 2737  
QY 2401 TCTGTTTTAGAACGTGAGCAGCATTCTTCTACCTCAATGCAACCCGACAAAACCTTCTAGAA 2460  
Db 2738 TCTGTTTTAGAACGTGAGCAGCATTCTTCTACCTCAATGCAACCCGACAAAACCTTCTAGAA 2797  
QY 2461 TATATCTGAGTTTTGTGGTGGCTG 2484  
Db 2798 TATATCTGAGTTTTGTGGTGGCTG 2821

## RESULT 6

ABX72233

ID ABX72233 standard; cDNA; 2684 BP.

XX AC ABX72233;

XX AC ABX72233;

DT 03-JUN-2003 (first entry)

XX Human NOVX polynucleotide #64.

DE DE

XX XX

XX XX

KW Human; NOVX; gene; ss; metabolic disorder; cardiomyopathy; diabetes; ASD;

KW hypertension; congenital heart defect; aortic stenosis; valve disease;

KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;



QY 841 GATGAACACCGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 900  
Db |||||  
QY 941 GATGAACACCGATGGACGGAAGATGTTCTCTATAGAGCTTTGCTCCATGGGCATATA 1000  
Db |||||  
QY 901 GTTGAATTTACTTCGACACGACCAATTTCTTGAGATCTTCTCGGAAGGCACACGTTCTAGG 960  
Db |||||  
QY 1001 GTTGAATTTACTTCGACACGACCAATTTCTTGAGATCTTCTCGGAAGGCACACGTTCTAGG 1060  
Db |||||  
QY 961 AGTGAAGAAACCTCTTTGCTCGGCAGGACTTTTGTGAGTTGTGGTAGACTCTGTCT 1020  
Db |||||  
QY 1061 AGTGAAGAAACCTCTTTGCTCGGCAGGACTTTTGTGAGTTGTGGTAGACTCTGTCT 1120  
Db |||||  
QY 1021 ACCAATGTCATCCAGACATCTTGATATATACCTGTTGGAATCTCTTATGATCGCATATC 1080  
Db |||||  
QY 1121 ACCAATGTCATCCAGACATCTTGATATATACCTGTTGGAATCTCTTATGATCGCATATC 1180  
Db |||||  
QY 1081 GAAGTCACTACAATGTTGAACCTGGGCAACCTTAAGAAATGAGAGCCTGTGGAGT 1140  
Db |||||  
QY 1181 GAAGTCACTACAATGTTGAACCTGGGCAACCTTAAGAAATGAGAGCCTGTGGAGT 1240  
Db |||||  
QY 1141 GTAGCAAGAGGTGTTATTAGAAATGTTACGAAACCTATGTTGTCCGAGTGGATTTT 1200  
Db |||||  
QY 1241 GTAGCAAGAGGTGTTATTAGAAATGTTACGAAACCTATGTTGTCTCCGAGTGGATTTT 1300  
Db |||||  
QY 1201 GCACAGCCATTTCTTTAAAGAAATTTTAAAGCCAAAGTCAGAAACCCGGTGTCTGCT 1260  
Db |||||  
QY 1301 GCACAGCCATTTCTTTAAAGAAATTTTAAAGCCAAAGTCAGAAACCCGGTGTCTGCT 1360  
Db |||||  
QY 1261 CTACTTTCCCTGGAGCAGCGTTGTTACGACTATCTTCTTCAAGACCCAGTATGCT 1320  
Db |||||  
QY 1361 CTACTTTCCCTGGAGCAGCGTTGTTACGACTATCTTCTTCAAGACCCAGTATGCT 1420  
Db |||||  
QY 1321 GCTGATGAAGGTAGAGACACGTCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1380  
Db |||||  
QY 1421 GCTGATGAAGGTAGAGACACGTCATTAATGAGTCCAGAAATGCAACAGATGAATCCCTA 1480  
Db |||||  
QY 1381 CGAAGGAGGTGATTGCAAAATCTGGCTGAGCATATTTCTATCTAGTCAAGTCCCTGT 1440  
Db |||||  
QY 1481 CGAAGGAGGTGATTGCAAAATCTGGCTGAGCATATTTCTATCTAGTCAAGTCCCTGT 1540  
Db |||||  
QY 1441 GCCATTATGTCACACACATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1500  
Db |||||  
QY 1541 GCCATTATGTCACACACATTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1600  
Db |||||  
QY 1501 GATCTCTCCATTTGGTTCGAAGACTTTCTTTGATGAAGAGGAAGTCTTGGCTCGTGAT 1560  
Db |||||  
QY 1601 GATCTCTCCATTTGGTTCGAAGACTTTCTTTGATGAAGAGGAAGTCTTGGCTCGTGAT 1660  
Db |||||  
QY 1561 TTTGACCTGGGGTTCTCAGGAAATTCAGAGATGTAGTAATGTCATGCCATACAGCTGCTG 1620  
Db |||||  
QY 1661 TTTGACCTGGGGTTCTCAGGAAATTCAGAGATGTAGTAATGTCATGCCATACAGCTGCTG 1720  
Db |||||  
QY 1621 GGAATTTGTTCACAAATCACCACACTAGCAGGAACGATGATGTTTTTATCACCACCCAGC 1680  
Db |||||  
QY 1721 GGAATTTGTTCACAAATCACCACACTAGCAGGAACGATGATGTTTTTATCACCACCCAGC 1780  
Db |||||  
QY 1681 ACAATGTCCTCAGTCTCGAATCAACTTCTACAGCAATGGGTACTTCAATGCTCTTT 1740  
Db |||||  
QY 1781 ACAATGTCCTCAGTCTCGAATCAACTTCTACAGCAATGGGTACTTCAATGCTCTTT 1840  
Db |||||  
QY 1741 ATCATGGAGGCCATCATAGCTTTGAGCCTTTTATGAGTCTTGAACAGAGGGGACTGGGG 1800  
Db |||||  
QY 1841 ATCATGGAGGCCATCATAGCTTTGAGCCTTTTATGAGTCTTGAACAGAGGGGACTGGGG 1900  
Db |||||  
QY 1801 GGTCCCACTAGCACCACCTAACTGATACGAGGAGCAGCTGGTGGGAAGCGCGCC 1860  
Db |||||  
QY 1901 GGTCCCACTAGCACCACCTAACTGATACGAGGAGCAGCTGGTGGGAAGCGCGCC 1960  
Db |||||  
QY 1861 AGCTGTGCTACCTTCTTCAATGAAGGACCATCTTCACTGCTTGGCAGACATTTTAC 1920  
Db |||||  
QY 1961 AGCTGTGCTACCTTCTTCAATGAAGGACCATCTTCACTGCTTGGCAGACATTTTAC 2020  
Db |||||

QY 1921 CAAGTCTGCCATGAAACAGTAGTAGGAAGTTTATCCAGTATGCGATTTCTTACAGTGGCAGAG 1980  
Db |||||  
QY 2021 CAAGTCTGCCATGAAACAGTAGTAGGAAGTTTATCCAGTATGCGATTTCTTACAGTGGCAGAG 2080  
Db |||||  
QY 1981 CACGATGACCCAGGAGATATCAGTCTCTAGTCTTCTGCTGAGCAGCAGTGGGCAAGAGCTT 2040  
Db |||||  
QY 2081 CACGATGACCCAGGAGATATCAGTCTCTAGTCTTCTGCTGAGCAGCAGTGGGCAAGAGCTT 2140  
Db |||||  
QY 2041 CCAGAACCTTTGCTTGGAGAAAGTGAAGAAAGTGAAGACAGTGAATTTGGGGAGGAA 2100  
Db |||||  
QY 2141 CCTGAACTTTGCTTGGAGAAAGTGAAGAAAGTGAAGACAGTGAATTTGGGGAGGAA 2200  
Db |||||  
QY 2101 CAGCGAGATTGCTACCTGAAGGTGAGCCAAATCCAAAGGAGCACCAGCAGTTTATCACCTTC 2160  
Db |||||  
QY 2201 CAGCGAGATTGCTACCTGAAGGTGAGCCAAATCCAAAGGAGCACCAGCAGTTTATCACCTTC 2260  
Db |||||  
QY 2161 TTACAGAGACTCTTGGGCTTTGCTGGAGCCCTACAGCTCTGCTGCCATCTTTGTTTAC 2220  
Db |||||  
QY 2261 TTACAGAGACTCTTGGGCTTTGCTGGAGCCCTACAGCTCTGCTGCCATCTTTGTTTAC 2320  
Db |||||  
QY 2221 AACTTTCAGTGGTCTCTGTTCCAGAACTGAGTATCTGCAAAAAGTTGCAAAAATACCTAATA 2280  
Db |||||  
QY 2321 AACTTTCAGTGGTCTCTGTTCCAGAACTGAGTATCTGCAAAAAGTTGCAAAAATACCTAATA 2380  
Db |||||  
QY 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTCAGAGTGCACATATGTTGTTGAG 2340  
Db |||||  
QY 2381 ACCAGAACAGAAAGAAATGTTGAGTATATGCTCAGAGTGCACATATGTTGTTGAG 2440  
Db |||||  
QY 2341 AATGCTGTGAAATGTTTAAAGGATATTTGGGGTTTTTCAAGGAGACCAAAAGAGAGTG 2400  
Db |||||  
QY 2441 AATGCTGTGAAATGTTTAAAGGATATTTGGGGTTTTTCAAGGAGACCAAAAGAGAGTG 2500  
Db |||||  
QY 2401 TCTGTTTTAGAACTGAGCAGCAGTCTTCTACCTCAATGCAACCGACCAAAACTTCTAGAA 2460  
Db |||||  
QY 2501 TCTGTTTTAGAACTGAGCAGCAGTCTTCTACCTCAATGCAACCGACCAAAACTTCTAGAA 2560  
Db |||||  
QY 2461 TATATCTGAGTTTTTGTGCTG 2484  
Db |||||  
QY 2561 TATATCTGAGTTTTTGTGCTG 2584  
Db |||||

## RESULT 7

ABZ12025  
ID ABZ12025 standard; cDNA; 3273 BP.  
XX AC ABZ12025;  
XX AC ABZ12025;  
XX DT 20-JAN-2003 (first entry)  
XX DE Human polynucleotide SEQ ID NO 907.  
XX DE Human; genome mapping; gene therapy; food supplement; virus; fungus;  
XX KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
XX KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
XX KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
XX KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;  
XX KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
XX KW haemostatic; vulnary; fungicide; antibacterial; virucide; protozoacide;  
XX KW antiarthritic; gene; ss.  
XX OS Homo sapiens.  
XX PN W0200270539-A2.  
XX XX 12-SEP-2002.  
XX PF 05-MAR-2002; 2002WO-US05095.  
XX PF 05-MAR-2001; 2001US-0799451.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Wehrman T, Wang J, Wang D, Drmanac RT;  
XX WPI: 2002-759812/82.  
DR P-PSDB; ABP69808.

XX New polynucleotides comprising sequences assembled from expressed  
PT sequence tags (ESTs), useful for treating cell-proliferative,  
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or  
PT platelet or coagulation disorders -

XX Claim 1; SEQ ID NO 907; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising a  
CC nucleotide sequence selected from any of 948 sequences  
CC (ABZ11119-ABZ12066) or their mature protein coding portion, active domain  
CC coding protein or complementary sequences. The polynucleotides are useful  
CC for identifying expressed genes or for physical mapping of human genome.  
CC The encoded polypeptides (ABP6902-ABP69849) are useful as molecular  
CC weight markers, as a food supplement, for generating antibodies, in  
CC medical imaging, screening and diagnostic assays and for treating  
CC cell-proliferative disorders (cancer), neurodegenerative diseases  
CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple  
CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid  
CC disorders, platelet or coagulation disorders, wound, burns, incision,  
CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,  
CC parasitic), arthritis, etc.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

XX Sequence 3273 BP; 915 A; 704 C; 733 G; 921 T; 0 other;

Query Match 88.0%; Score 2185.8; DB 24; Length 3273;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2187; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 296 CAAGACACCGGATGGCTTGCAGACGCTTCTTACGTTCTTTTATTCAAGACGAG 355  
DB 173 CCAGACACCGGATGGCTTGCAGACGCTTCTTACGTTCTTTTATTCAAGACGAG 232  
QY 356 ATGTGCAATAGGCGCATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAG 415  
DB 233 ATGTGCAATAGGCGCATGTTTGGCCACCAATGTGACTGAAATGTGCTGAACAGCAGTAGAG 292  
QY 416 TACAAGAGCAATTCAGAAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCAGCAGC 475  
DB 293 TACAAGAGCAATTCAGAAAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCAGCAGC 352  
QY 476 AATCAAAAGCCGTTAACAAGTGAAAGAGCTAAAGAGTCTTCTCAAGAAATGGTTG 535  
DB 353 AATCAAAAGCCGTTAACAAGTGAAAGAGCTAAAGAGTCTTCTCAAGAAATGGTTG 412  
QY 536 CCACTGTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTTAAACCTGTTCAACA 595  
DB 413 CCACTGTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGCTTAAACCTGTTCAACA 472  
QY 596 GCTTCTTTTGGAAACATTTCAAAATTCAGAAAGTCACTTGAGATGGTTAAAGCTGCAACTG 655  
DB 473 GCTTCTTTTGGAAACATTTCAAAATTCAGAAAGTCACTTGAGATGGTTAAAGCTGCAACTG 532  
QY 656 AGACGAATTTGGCGTCTCTGTTTCTACAGTTTCATAGATCCCATATTTGACTATCTGCTGC 715  
DB 533 AGACGAATTTGGCGTCTCTGTTTCTACAGTTTCATAGATCCCATATTTGACTATCTGCTGC 592  
QY 716 TCACCTTTTCACTTCTCTGCCATAACATCAAAAGCACCATACATTTGCTTCAGGCAATAATC 775  
DB 593 TCACCTTTTCACTTCTCTGCCATAACATCAAAAGCACCATACATTTGCTTCAGGCAATAATC 652  
QY 776 TCACCATCCCAATCTTCAGTACCTTGATCCGTAAGCTTGGGGCTTCTTCTACAGCAGAA 835  
DB 653 TCAACATCCCAATCTTCAGTACCTTGATCCGTAAGCTTGGGGCTTCTTCTACAGCAGAA 712

QY 836 GGCTCGATGAAACACCCAGATGACCGGAAAGATGTTTCTTATAGAGCTTTGCTCCATGGGC 895  
DB 713 GGCTCGATGAAACACCCAGATGACCGGAAAGATGTTTCTTATAGAGCTTTGCTCCATGGGC 772  
QY 896 ATATAGTTGAATTAATCTTGCAGAGCAATTTCTTGAGATCTTCTTGGAAAGCAGCAGTT 955  
DB 773 ATATAGTTGAATTAATCTTGCAGAGCAATTTCTTGAGATCTTCTTGGAAAGCAGCAGTT 832  
QY 956 CTAGAGATGGAAAAACCTTCTGCTCGGGCAGGACTTTTGTGAGTTGTTGGTAGATACTC 1015  
DB 833 CTAGAGATGGAAAAACCTTCTGCTCGGGCAGGACTTTTGTGAGTTGTTGGTAGATACTC 892  
QY 1016 TGTCTACCAATGCTATCCAGACATCTTGTATATACCTGTTGGAAATCTCTTATGATCGCA 1075  
DB 893 TGTCTACCAATGCTATCCAGACATCTTGTATATACCTGTTGGAAATCTCTTATGATCGCA 952  
QY 1076 TTATCGAAGGTCACTACAAATGTTGAACAACTGGGCAACCTAAGAAAGATGAGAGCTGT 1135  
DB 953 TTATCGAAGGTCACTACAAATGTTGAACAACTGGGCAACCTAAGAAAGATGAGAGCTGT 1012  
QY 1136 GGAGTGTAGCAAGAGGTGTTATTAGAACTTTACGAAAAAACTATGTTGTTCCGAGTGG 1195  
DB 1013 GGAGTGTAGCAAGAGGTGTTATTAGAACTTTACGAAAAAACTATGTTGTTCCGAGTGG 1072  
QY 1196 ATTTTGCAGAGCAATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGT 1255  
DB 1073 ATTTTGCAGAGCAATTTTCTTAAAGGAATATTTAGAAAGCCAAAGTCAGAAACCGGTGT 1132  
QY 1256 CTGCTCTACTTCTTCCCTGGAGCAAGGTGTTACCAAGCTATATCTTCTTCAAGACCCAGTG 1315  
DB 1133 CTGCTCTACTTCTTCCCTGGAGCAAGGTGTTTACCAGCTATATCTTCTTCAAGACCCAGTG 1192  
QY 1316 ATGCTGTGATGAAGGTAGAGACAGCTCCATTAATAGTCCAGAAATCAACAGATGAAT 1375  
DB 1193 ATGCTGTGATGAAGGTAGAGACAGCTCCATTAATAGTCCAGAAATCAACAGATGAAT 1252  
QY 1376 CCCTACGAAGGAGGTGATTTGCAAAATCTGGCTGAGCATATTTCTTACTGCTAGCAAGT 1435  
DB 1253 CCCTACGAAGGAGGTGATTTGCAAAATCTGGCTGAGCATATTTCTTACTGCTAGCAAGT 1312  
QY 1436 CTGTGCAATTAATGTTCCACACATTTGTTGGCTTGGCTTCTCTTACAGACACAGCAGG 1495  
DB 1313 CCTGTGCCATTAATGTTCCACACATTTGTTGGCTTGGCTTCTCTTACAGACACAGCAGG 1372  
QY 1496 GAATTTGATCTCTCCACATTTGTTGGAAGCTTCTTGTGATGAAGAGGAGTCTTGGCTC 1555  
DB 1373 GAATTTGATCTCTCCACATTTGTTGGAAGCTTCTTGTGATGAAGAGGAGTCTTGGCTC 1432  
QY 1556 GTGATTTTGAACCTGGGTTCTCAGGAAATTCAGAAAGTGTAGTAATGCATGCCATACAG 1615  
DB 1433 GTGATTTTGAACCTGGGTTCTCAGGAAATTCAGAAAGTGTAGTAATGCATGCCATACAG 1492  
QY 1616 TGCTGGGAAATTTGTGCAAAATCAACCACTAGCAGGAAACGATGAGTTTTTATCACCC 1675  
DB 1493 TGCTGGGAAATTTGTGCAAAATCAACCACTAGCAGGAAACGATGAGTTTTTATCACCC 1552  
QY 1676 CCAGCACAACCTGCTCCATCAGTCTTGAACCTCAACTTCTTACAGCAATGGGTACTTTCATG 1735  
DB 1553 CCAGCACAACCTGCTCCATCAGTCTTGAACCTCAACTTCTTACAGCAATGGGTACTTTCATG 1612  
QY 1736 TCTTTATCATGAGGCGCATATAGCTTGCAGGCTTTATGCAAGTCTTGAACAGAGGGGAC 1795  
DB 1613 TCTTTATCATGAGGCGCATATAGCTTGCAGGCTTTATGCAAGTCTTGAACAGAGGGGAC 1672  
QY 1796 TGGGGGTTCCCACTAGCAGCCCACTTAACCTGATCAGCAGGAGCAGCTGTTGCGGAAGG 1855  
DB 1673 TGGGGGTTCCCACTAGCAGCCCACTTAACCTGATCAGCAGGAGCAGCTGTTGCGGAAGG 1732  
QY 1856 CGGCCAGCTGCTGCTTCTCTCCATGAGGACCACTTCTCAGTCTGCTGCTTCCAGACAT 1915  
DB 1733 CGGCCAGCTGCTGCTTCTCTCCATGAGGACCACTTCTCAGTCTGCTGCTTCCAGACAT 1792  
QY 1916 TTTACCAAGTCTGCGCATGAAACAGTAGGAAAGTTTTATCCAGTATGGCATTTCTTACAGTGG 1975



Db 520 CACAAGGCGATGTTTCCACCACTATCACTGACAAATGTACTGAATAGCAGCAGAGTCCAA 579  
Qy 421 GAGGCAATTCAGAGTGGCTGCTGAATTAACCCCTGATGGTTCTGCCAGCAGCAATCA 480  
Db 580 GAGGCAATTCGCTGAGTGGCTGAGAAATGAACCCGGATGGATCTGCCAGCAGAGTCC 639  
Qy 481 AAAGCGGTTAAACAAAGTGAAGAAAGAGCTTAAAGGATCTTCAAGAAATGGTTGCCACT 540  
Db 640 AAAGCCATCCAGAAATGAAGAAAGAGCCAGGAAGATCCTCCAGGAATGGTTGTACA 699  
Qy 541 GTCTCACCGGCAATGATCAGACTGACTGGGTGGGTGCTGATAAACTGTTCAACAGCTTC 600  
Db 700 GTCTCCCGGGATGATCAGGCTGACTGGCTGGGTGTTACTAAAGCTCTTCAACAGCTTC 759  
Qy 601 TTTTGGAAACATTCABAATTCABAAGGTCACATTTGAGATGGTTAAAGCTCCACTGAGAG 660  
Db 760 TTTCTGGAACATTCAGATTCACAAGGTCACATTTGAGATGGTTAAAGCTCCACTGAGAG 819  
Qy 661 AATTTGCCGCTTCTGTTTCTACCAAGTTCATAGATCCCATATTTGACTATCTGTGCTCACT 720  
Db 820 AATCTGCCGCTTCTGTTTCTGCCGCTGCACAGATCCACATCGACTACCTGCTGCTCACC 879  
Qy 721 TTCATTTCTCTTCCGCAATCAACAAAGCAACCATACATTTGCTTTCAGGCAATTAATCTCAAC 780  
Db 880 TTCATCTCTTCTGCCACAACATCAAAAGCTCCATACATCGCTCGGGCAACAACTCAAC 939  
Qy 781 ATCCCAATCTTCAGTACCTTGATCCATTAAGCTTGGGGCTTCTTCAACAGCAGAGCTTC 840  
Db 940 ATCCCAATCTTCAGTACCTTGATTCACAAGCTTGGGGCTTCTTCAACAGCAGAGCTTC 999  
Qy 841 GATGAACACCAAGATGGAAGAGATGTTCTCTATAGAGCTTGTCTCATGGGCATATA 900  
Db 1000 GACGAACCTCCAGATGGAAGAGATGTTCTCTATAGAGCTTGTCTCATGGGCATATA 1059  
Qy 901 GTTGAATTAATTCGACAGCAGCAATTTCTGGAGATCTTCTGGAAGGCAACGTTCTAGG 960  
Db 1060 GTTGAATCTCTCCGACAGCAGCAGTCTCTGGAGATCTTCTGGAAGGCAACCGCTCCCG 1119  
Qy 961 AGTGGAAACCTTCTGTCTCGGCAGGACTTCTGTGATGTTGTGATGATCTGTCT 1020  
Db 1120 AGTGGCAAGACCTCTGTGTCCCGGGCGGGCTCTGTGATGTTGTGATGATCTGTCTCA 1179  
Qy 1021 ACCAATGTCTATCCAGACATCTTGAATAATACCTGTTGGAATCTCTATGATCGCATATC 1080  
Db 1180 TCCNAACCATCTCTGACATCTCTGTTCTATCCCTGTGGGATCTCTGATGATCGATAATC 1239  
Qy 1081 GAAGGTCACTACAATGGTGAACAACTGGGCAACCTTAAGAAGATGAGAGCCTGTGGAT 1140  
Db 1240 GAAGGTCACTACAATGGTGAACAACTGGGCAACCTTAAGAAGATGAAAGTCTCTGGAT 1299  
Qy 1141 GTAGCAAGGTTGTTATTAAGATGTTACGNAANAATATGTTGTGTTGTTGTTGTTGTT 1200  
Db 1300 GTGGCAAGAGGCTTATCAAGATGTTGCGGAAATCTAGCGCTATCTGCGAGTGGACTTT 1359  
Qy 1201 GCACAGCCTTTTCTTAAAGGAATTTAGAAAGCCAAAGTCAGAAAACCGGTGTCTGCT 1260  
Db 1360 GCACAGCCTTTTCTTAAAGGAATTTAGAAAGCCAAAGTCAGAAAACCTGTATCTGCT 1419  
Qy 1261 CTACTTTCCCTGGAGCAAGGTTTGTACCAGCTATATCTTCTTCAAGACCCAGTATGCT 1320  
Db 1420 CCGCTCTCTTTGGAGCAAGCAGTGTACCAGCAATCTTCTTCAAGACCTGTATGCTGCT 1479  
Qy 1321 GCTGATGAAGGTAGACACAGTCCATTAATGATCCAGAAATGCAACAGATGATCCCTTA 1380  
Db 1480 GCTGCCGAAATGAAGACATGTCCAGTAATGATGAGTCCAGAAACCGGCAAGCAAGCTTC 1539  
Qy 1381 CGAAGGAGGTTGATTCGAAATCTGGCTGAGCATATTTCTATTCTACTGCTAGCAAGTCTCT 1440  
Db 1540 CGAAGGAGGCTGATCGAAGACCTTGGCGGAGCAGTCTCTTCAACCGCAGCAAGTCTCT 1599  
Qy 1441 GCCATTTATGTCCACACATTTGGCTTGGCTGCTCTCTTACAGACACAGGCGGGAAT 1500

Db 1600 GCTATCATGTCCACCCACATTTGTGGCTCTCTCTACAGACACAGGAGGGAATC 1659  
Qy 1501 GATCTCTCCACATTTGTGGAAGACTTCTTTGTGATGAAGAGAAAGTCTGCTGCTGTAT 1560  
Db 1660 CACCTCTCCACGCTGTGGAAGACTTCTTTGTGATGAAGAGAAAGTCTTACTGCTGGAT 1719  
Qy 1561 TTTGACCTGGGTTCTCAGGAAATTCAGAAGATGTAGTAATGCATGCCATACAGCTGCTG 1620  
Db 1720 TTTGACCTGGGTTCTCAGGAAATTCAGAAGATGTAGTCATGCTATTCAGCTTCTG 1779  
Qy 1621 GGAATTTGTGACAACTACCCACACTAGCAGGAAGAGTATTTTATCACCACCCAGC 1680  
Db 1780 GGAATTTGTGACAACTACCCACACTAGCAGGAAGAGTATTTTATTTACTCCAGC 1839  
Qy 1681 ACAATGTCTCCATCAGTCTTCGAACTCAACTTCTACAGCAATGGGTACTTCTATGCTTT 1740  
Db 1840 ACAATGTCTCCGCTCGTCTTTGAACTCAACTTCTACAGCAATGGGTACTTCTATGCTTT 1899  
Qy 1741 ATCATGGAGGCGCATCATAGCTTTGAGCCCTTTATGAGCTTCTGAACAAGAGGAGCTGGG 1800  
Db 1900 ATCATGGAGGCGCATCATAGCTTTGAGCAATTTATGAGTCCAGAAATGAAGAGGCTTCCGGA 1959  
Qy 1801 GGTCCCATAGCACCACCTTAACCTGATCAGCAGGAGCAGCTGGTGGGAAGGGGCGC 1860  
Db 1960 GGTCTGCGGAGGCGCTTGGCAACCTGATCAGCAGGAGCAGCTGGTGGGAAGGGCGCGC 2019  
Qy 1861 AGCCTGTGCTACTTCTCTCCAAATGAAGCACCCTCTCACTGCTTGGCCAGACATTTTAC 1920  
Db 2020 AGCCTGTGCTACTTCTCTTAATGAAGTACATTTCTCTGCCCTGCCAGACATTTTAC 2079  
Qy 1921 CAAAGTCTGCATGAACAGTAGGAAAGTTTATCCAGTATGGCAATTTTACAGTGGCAGAG 1980  
Db 2080 CAGGTTTGTGAAGACAGTAGGAAAGTTTATCCAGTACGGAATTTCTCACAGTGGCAGAG 2139  
Qy 1981 CAGATGACAGGAAGATATCAGTCTTAGTCTTGTGAGCAGCAGTGGGCAAGAGCTT 2040  
Db 2140 CAAAGTACAGGAAGATGTCTAGTCTGCGCTTGCAGAGCAGCAGTGGAAACAGAGCTT 2199  
Qy 2041 CCAGAACCTTTCTCTTGGAGAGTGTGAAGAGATGAAGACAGTGTCTTTCGGGAGGAA 2100  
Db 2200 CCGAGCCTCTGAACTGGAGAAAGTGCAGAAAGATGAGGACAGTGTCTTGTGAGGAG 2259  
Qy 2101 CAGCAGATTTGCTTACTGAAAGTGAAGCAATCCAAAGGAGCACCAGCAGTTTATCACCCTTC 2160  
Db 2260 CAGCCTGATTTGCTACTCTGAAGGTGAGCCAGGCCAAGGAGCACCAGCAATTTATCACCCTTT 2319  
Qy 2161 TTACAGAGACTCTTGTGGGCTTTGTGGAGGCTTACAGCTGTGCTGCCATCTTTGTTTAC 2220  
Db 2320 CTGCAAGGCTTCTGGGGCCCTGTCTAGAAAGCTTACAGCTGTGCTGCCATCTTTGTCCAC 2379  
Qy 2221 AACTTCAGTGTCTTCTCCAGAACCTGAGTATCTGAAAGTTCACAAATACCTTAATA 2280  
Db 2380 ACTTTCGGGGCCAGTCCCGAGTCTGAGTACCTGCAAGAGTGCACAGGTACCTTCTC 2439  
Qy 2281 ACCAGAACAGAAAGAAATGTTGAGTATATGCTGAGAGTGCCACATATTTGTTGTGAAG 2340  
Db 2440 ACCAGAACAGGAGGAGCAGTCCGCGGTGATGCTGAGAGTGCCACATCTGTCTGTGAAG 2499  
Qy 2341 AATGCTGGAATAATGTTTAAAGATATTTGGGTTTTTCAAGGAGACCAACAAAGAGAGTG 2400  
Db 2500 AATGCTGGAATAATGTTTAAAGATATTTGGGTTTTTCAAGGAGACCAAGCAGAGGAGCG 2559  
Qy 2401 TCTGTTTTTAGACTCAGCAGCAGCTTTTCTTACCTCAATGCAACCGCAAAACCTTCTAGAA 2460  
Db 2560 TCTGTTTTTAGAACTCAGCAGCAGCTTTTCTTACCTCAGGCGAGCGGAGAGCTTCTTGGAA 2619  
Qy 2461 TATATTCTGAGTTTTTGTGGTGTGCTG 2484  
Db 2620 TACATTCTGAGCTTCTGTGTGTGCTG 2643









CC	seqdata.uspto.gov/sequence.html?DocID=20020137139.	PT	analysis, cattle breeding, or for genetically improving cattle -
XX	Sequence 407 BP; 108 A; 101 C; 89 G; 109 T; 0 other;	PS	Claim 2; SEQ ID No 742; 245pp; English.
XX	Query Match 13.6%; Score 337.6; DB 25; Length 407;	XX	The invention relates to a purified nucleic acid molecule associated with
XX	Best Local Similarity 90.9%; Pred. No. 1.1e-94;	CC	lactation or muscle and fat deposition (designated LMPD), derived
XX	Matches 371; Conservative 0; Mismatches 34; Indels 3; Gaps 1;	CC	from cattle, and the LMPD nucleic acid can specifically hybridise to a
QY	1330 GGTAGACACGTCCTTAAATGAGTCCAGAAATGCAACAGATGAATCCCTACGAAGGAGG 1389	CC	second nucleic acid molecule comprising any of 15112 nucleotide
Db	3 GATACAGACAAGTCCATTAATGAGTCCAGAAATGCAACAGATGAATCCCGCGCA---AGA 59	CC	sequences, appearing as ABX34836-ABX49947, or complements of them.
QY	1390 TTGATTGCAAACTGGCTGAGCATATTTCTTACTGTAGCAAGTCTGTGCCATTATG 1449	CC	Also included are: (1) a transformed cell having a nucleic acid
Db	60 CTGATTGCACATCTGGCTGAGCACATTTCTTCTACTGTAGCAAGTCTGTGCTATTATG 119	CC	comprising an LMPD nucleic acid linked to a promoter and a 3' non-
QY	1450 TCCACACATTTGGCTTGGCTTGGCTCTCTCTACAGACAGCAGGGAATTCATCTCC 1509	CC	translated sequence that functions in the cell to cause termination of
Db	120 TCCAGCACATCTGGCTTGGCTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 179	CC	transcription and addition of polyadenylated ribonucleotides to a 3' end
QY	1510 ACATTGGTGAAGACTCTTTTGTGATGAAGAGGAGTCTTGGCTCGTGAATTTGACCTG 1569	CC	of the mRNA molecule; and (2) determining a level or pattern of a
Db	180 ACATTGGTGAAGACTCTTTTGTGATGAAGAGGAGTCTTGGCTCGTGAATTTGACCTG 239	CC	molecule in a bovine cell or tissue comprising: (a) incubating a marker
QY	1570 GGGTCTCAGGAAATTCAGAAATGATAGTAAATGATGCGATGACAGTCTGTGGAAATGT 1629	CC	nucleic acid (comprising any of the 15112 nucleic acid sequences or its
Db	240 GGGTCTCAGGAAATTCAGAAATGATAGTAAATGATGCGATGACAGTCTGTGGAAATGT 299	CC	complement or fragment) with a complementary nucleic acid molecule
QY	1630 GTCAAAATCACCACACTAGCAGGAGCATGATGATTTTATCACCCTCCAGCACACTGTC 1689	CC	obtained from the bovine cell or tissue, where hybridisation between the
Db	300 ATCAAAATCACCACACTAGCAGGAGCATGATGATTTTATCACCCTCCAGCACACTGTC 359	CC	marker nucleic acid and the complementary nucleic acid permits the
QY	1690 CCATCAGTCTTGGAACTCAACTTCTACAGCAATGGGGTACTTCATGTC 1737	CC	detection of the molecule; and (b) detecting the level or pattern of the
Db	360 CCATCAGTCTTGGAACTCAACTTCTACAGCAATGGGGTACTTCATGTC 407	CC	complementary nucleic acid, where the detection of the complementary
RESULT 12		CC	nucleic acid is predictive of the level or pattern of the molecule.
ID	ABX35577/c	CC	The LMPD nucleic acid is used for determining a level or pattern
XX	ABX35577 standard; cDNA; 418 BP.	CC	of a molecule in a bovine cell or tissue. It is useful for genome
XX	AC ABX35577;	CC	mapping, gene identification and analysis, cattle breeding, preparation
XX	20-FEB-2003 (first entry)	CC	of constructs for use in cattle gene expression, or for genetically
DE	Bovine EST associated with lactation/muscle/fat deposition #742.	CC	improving cattle. The present sequence is one of the 15112 bovine
XX		CC	LMPD EST (expressed sequence tag) nucleic acids.
KW	Bovine; ss; EST; expressed sequence tag; lactation; LMPD;	CC	Note: The present sequence was not shown in the specification but
KW	muscle deposition; fat deposition; genome mapping; gene identification;	CC	was obtained in electronic format from the USPTO web site:
KW	gene analysis; cattle breeding.	CC	seqdata.uspto.gov/sequence.html?DocID=20020137139.
XX	Bos Taurus.	SQ	Sequence 418 BP; 97 A; 116 C; 102 G; 102 T; 1 other;
OS		Query Match 12.2%; Score 302; DB 25; Length 418;	
XX	US2002137139-A1.	Best Local Similarity 84.1%; Pred. No. 1.6e-83;	
XX	26-SEP-2002.	Matches 354; Conservative 0; Mismatches 61; Indels 6; Gaps 1;	
XX	24-SEP-2001; 2001US-0960352.	QY	1772 ATGCAGTTCTCAACAAGAGGGGACTGGGGGTGCCACTAGCACCCACCTAACCTGATCA 1831
XX	12-JAN-1999; 99US-115707P.	Db	418 ATGCAGTTCTCAACAAGAGGGGCTCCAGGAGCC-----CCGTCCTCCAGCTTGGTCA 365
PR	11-JAN-2000; 2000US-0480902.	QY	1832 GCCAGGACAGCTGGTGGGAAAGGCGGCGCAGCTGTGCTACTCTTCTTCCCAATGAAGGCA 1891
XX	(BYAT/) BYATT J C.	Db	364 GCCAGGACAGCTGGTGGCACAAGGCTGCCAGGGTGTGCTATCTGCTCTCCCAATGAAGGCA 305
PA	(MATH/) MATHIALAGAN N.	QY	1892 CCATCTCACTGCTTGGCAGACATTTTACCAAGTCTGCCATGAAACAGTGAAGATTTA 1951
PA	(TAON/) TAO N.	Db	304 CCATCTCTCTCCCTGCCAGACCTTTTACCCTAAATTTGCCATGAACAGTGGCGGTTTA 245
PA	(WARR/) WARREN W C.	QY	1952 TCCAGTATGGCATTTCTTACAGTGGCAGACAGATGACGAGAGATATCAGTCTAGTC 2011
XX	Byatt JC, Mathialagan N, Tao N, Warren WC;	Db	244 TCCAGTATGGCAATTTTATAGTTCCGAGCAAGATAATCAGGAAGATATCAGTCTTGT 185
XX	WPI; 2003-110599/10.	QY	2012 TTGCTGAGCAGCAGTGGGACAAAGAGCTTCAGAAACCTTTCTTGGAGAGTGTATGAAG 2071
XX	New nucleic acid associated with lactation, and muscle and fat	Db	184 CTTCGAGCAGCAGTGGGACAAAGAGTTCGGAACCTTTCTTGGAGAGTGTATGAAG 125
PT	deposition, useful for genome mapping, gene identification and	QY	2072 AAGATGAAGACAGTGAATTTGGGGAGGAAACAGCAGATTTGCTACTGAAGGTGAGCCAA 2131

ID ABX47921 standard; cDNA; 303 BP.  
XX AC ABX47921;  
DT 21-FEB-2003 (first entry)  
XX  
DE Bovine EST associated with lactation/muscle/fat deposition #13086.  
XX  
XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
KW muscle deposition; fat deposition; genome mapping; gene identification;  
KW gene analysis; cattle breeding.  
XX  
XX Bos Taurus.  
XX  
XX US2002137139-A1.  
XX  
XX 26-SEP-2002.  
XX  
XX 24-SEP-2001; 2001US-0960352.  
XX  
XX 12-JAN-1999; 99US-115707P.  
PR 11-JAN-2000; 2000US-0480902.  
XX  
PA (BYAT/) BYATT J C.  
PA (MATH/) MATHIALAGAN N.  
PA (TAON/) TAO N.  
PA (WARR/) WARREN W C.  
XX  
PI Byatt JC, Mathialagan N, Tao N, Warren WC;  
XX  
XX WPI; 2003-110599/10.  
XX  
XX New nucleic acid associated with lactation, and muscle and fat  
PT deposition, useful for genome mapping, gene identification and  
PT analysis, cattle breeding, or for genetically improving cattle  
XX  
XX Claim 2; SEQ ID No 13086; 245pp; English.  
XX  
XX The invention relates to a purified nucleic acid molecule associated with  
CC lactation or muscle and fat deposition (designated LMFD), derived to a  
CC from cattle, and the LMFD nucleic acid can specifically hybridise to a  
CC second nucleic acid molecule comprising any of 1512 nucleotide  
CC sequences, appearing as ABX34836-ABX49947, or complements of them.  
CC Also included are; (1) a transformed cell having a nucleic acid  
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-  
CC translated sequence that functions in the cell to cause termination of  
CC transcription and addition of polyadenylated ribonucleotides to a 3' end  
CC of the mRNA molecule; and (2) determining a level or pattern of a  
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker  
CC nucleic acid (comprising any of the 1512 nucleic acid sequences or its  
CC complement or fragment) with a complementary nucleic acid molecule  
CC obtained from the bovine cell or tissue, where hybridisation between the  
CC marker nucleic acid and the complementary nucleic acid permits the  
CC detection of the molecule; and (b) detecting the level or pattern of the  
CC complementary nucleic acid, where the detection of the complementary  
CC nucleic acid is predictive of the level or pattern of the molecule.  
CC The LMFD nucleic acid is used for determining a level or pattern  
CC of a molecule in a bovine cell or tissue. It is useful for genome  
CC mapping, gene identification and analysis, cattle breeding, preparation  
CC of constructs for use in cattle gene expression, or for genetically  
CC improving cattle. The present sequence is one of the 1512 bovine  
CC LMFD EST (expressed sequence tag) nucleic acids.  
CC Note: The present sequence was not shown in the specification but  
CC was obtained in electronic format from the USPTO web site:  
CC seqdata.uspto.gov/sequence.html?DocID=20020137139.  
XX  
XX Sequence 303 BP; 66 A; 90 C; 64 G; 83 T; 0 other;

Query Match 9.4%; Score 233; DB 25; Length 303;  
Best Local Similarity 87.9%; Pred. No. 5.3e-62;  
Matches 254; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
QY 1907 GCCAGACATTTTACCAAGTCTGCCATGAAACAGTAGGAAAGTTTATCCAGTAGGCATTC 1966

Db 303 GCCAGACGTTTATCCCAAATTTGCCATGAACAGTGGCGAGGTTTATCCAGTAGGCATTC 244  
QY 1967 TTACAGTGGCAGGACGACGATGACCCAGGAAGATATCAGTCTCTAGTCTTGTCTGACGACAGT 2026  
Db 243 TTATAGTGGCGAGCAAGATGATCAGGAAGATATCAGTCTGTAGTCTTGTCCGACGACAGT 184  
QY 2027 GGGACAAGAAGCTTCCAGAACCTTTGCTTTGGAGAAGTATGATGAAGAAGATGAAGACAGTG 2086  
Db 183 GGGACAAGAAGCTTCCGGAACATTTGTCTTTGGAGAAGTATGATGAAGAAGATGATGACAGCG 124  
QY 2087 ACTTTGGGAGGAGAAACAGCCAGATTTGCTACTGAAGTGAAGCCATTCGAAGGAGCACCAGC 2146  
Db 123 ATTATGCTGAGGAGCAACGAGATTGATACATGAAGGTGAGCCAGTCCAAGGAGCACCAGC 64  
QY 2147 AGTTTATCACCTTCTTACAGAGACTCTTTGGGCTTTGCTGGAGGCGCTA 2195  
Db 63 AGTTTATCACATTTCTCTGCAGAGGCTCTCGGACCTCTGCTGGAGGCGTA 15  
RESULT 14  
ABX41074  
ID ABX41074 standard; cDNA; 264 BP.  
XX  
XX AC ABX41074;  
XX  
XX 20-FEB-2003 (first entry)  
XX  
XX Bovine EST associated with lactation/muscle/fat deposition #6239.  
XX  
XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
KW muscle deposition; fat deposition; genome mapping; gene identification;  
KW gene analysis; cattle breeding.  
XX  
XX Bos Taurus.  
XX  
XX US2002137139-A1.  
XX  
XX 26-SEP-2002.  
XX  
XX 24-SEP-2001; 2001US-0960352.  
XX  
XX 12-JAN-1999; 99US-115707P.  
PR 11-JAN-2000; 2000US-0480902.  
XX  
XX (BYAT/) BYATT J C.  
PA (MATH/) MATHIALAGAN N.  
PA (TAON/) TAO N.  
PA (WARR/) WARREN W C.  
XX  
PI Byatt JC, Mathialagan N, Tao N, Warren WC;  
XX  
XX WPI; 2003-110599/10.  
XX  
XX New nucleic acid associated with lactation, and muscle and fat  
PT deposition, useful for genome mapping, gene identification and  
PT analysis, cattle breeding, or for genetically improving cattle  
XX  
XX Claim 2; SEQ ID No 6239; 245pp; English.  
XX  
XX The invention relates to a purified nucleic acid molecule associated with  
CC lactation or muscle and fat deposition (designated LMFD), derived  
CC from cattle, and the LMFD nucleic acid can specifically hybridise to a  
CC second nucleic acid molecule comprising any of 1512 nucleotide  
CC sequences, appearing as ABX34836-ABX49947, or complements of them.  
CC Also included are; (1) a transformed cell having a nucleic acid  
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-  
CC translated sequence that functions in the cell to cause termination of  
CC transcription and addition of polyadenylated ribonucleotides to a 3' end  
CC of the mRNA molecule; and (2) determining a level or pattern of a  
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker  
CC nucleic acid (comprising any of the 1512 nucleic acid sequences or its  
CC complement or fragment) with a complementary nucleic acid molecule  
CC obtained from the bovine cell or tissue, where hybridisation between the  
CC marker nucleic acid and the complementary nucleic acid permits the  
CC detection of the molecule; and (b) detecting the level or pattern of the  
CC complementary nucleic acid, where the detection of the complementary  
CC nucleic acid is predictive of the level or pattern of the molecule.  
CC The LMFD nucleic acid is used for determining a level or pattern  
CC of a molecule in a bovine cell or tissue. It is useful for genome  
CC mapping, gene identification and analysis, cattle breeding, preparation  
CC of constructs for use in cattle gene expression, or for genetically  
CC improving cattle. The present sequence is one of the 1512 bovine  
CC LMFD EST (expressed sequence tag) nucleic acids.  
CC Note: The present sequence was not shown in the specification but  
CC was obtained in electronic format from the USPTO web site:  
CC seqdata.uspto.gov/sequence.html?DocID=20020137139.  
XX  
XX Sequence 303 BP; 66 A; 90 C; 64 G; 83 T; 0 other;



**THIS PAGE BLANK (USPTO)**